

Second Order Optimality Conditions for Optimal Control Problems on Riemannian Manifolds*

Qing Cui[†] Li Deng[‡] and Xu Zhang[§]

Abstract This work is concerned with an optimal control problem on a Riemannian manifold, for which two typical cases are considered. The first case is when the endpoint is free. For this case, the control set is assumed to be a separable metric space. By introducing suitable dual equations, which depend on the curvature tensor of the manifold, we establish the second order necessary and sufficient optimality conditions of integral form. In particular, when the control set is a Polish space, the second order necessary condition is reduced to a pointwise form. As a key preliminary result and also an interesting byproduct, we derive a geometric lemma, which may have some independent interest. The second case is when the endpoint is fixed. For this more difficult case, the control set is assumed to be open in an Euclidian space. We obtain the second order necessary and sufficient optimality conditions, in which the curvature tensor also appears explicitly. Our optimality conditions can be used to recover the following famous geometry result: Any geodesic connecting two fixed points on a Riemannian manifold satisfies the second variation of energy; while the existing optimality conditions in control literatures fail to give the same result.

Keywords Optimal control, second order necessary and sufficient conditions, Riemannian manifold, curvature tensor

MSC (2010) 49K15, 49K30, 93C15, 58E25, 70Q05

*This work is supported by the National Science Foundation of China under grants 11401491 and 11231007, the Fundamental research funds for the Central Universities under grant 2682014CX052, and the grant MTM2014-52347 from the Spanish Science and Innovation Ministry.

[†]School of Mathematics, Southwest Jiaotong University, Chengdu 611756, Sichuan Province, China.
E-mail: cuiqing@home.swjtu.edu.cn.

[‡]School of Mathematics, Southwest Jiaotong University, Chengdu 611756, Sichuan Province, China.
E-mail: dengli@home.swjtu.edu.cn.

[§]School of Mathematics, Sichuan University, Chengdu 610064, Sichuan Province, China. *E-mail:* zhang_xu@scu.edu.cn.

1 Introduction

Let $n \in \mathbb{N}$ and M be a complete simply connected, n -dimensional manifold with Riemannian metric g . Let ∇ be the Levi-Civita connection on M related to g , $\rho(\cdot, \cdot)$ be the distance function on M , $T_x M$ be the tangent space of M at $x \in M$, and $T_x^* M$ be the cotangent space. Denote by $\langle \cdot, \cdot \rangle$ and $|\cdot|$ the inner product and the norm over $T_x M$ related to g , respectively. Also, denote by $TM \equiv \bigcup_{x \in M} T_x M$, $T^*M \equiv \bigcup_{x \in M} T_x^* M$ and $C^\infty(M)$ the tangent bundle, the cotangent bundle and the set of smooth functions on M , respectively.

Let $T > 0$, U be a metric space, and $f : [0, T] \times M \times U \rightarrow TM$ and $f^0 : [0, T] \times M \times U \rightarrow \mathbb{R}$ be two functions (satisfying suitable assumptions to be given later). Given $y_0 \in M$, let us consider the following control system

$$\begin{cases} \dot{y}(t) = f(t, y(t), u(t)), & a.e. t \in [0, T], \\ y(0) = y_0, \end{cases} \quad (1.1)$$

where $\dot{y}(t) = \frac{d}{dt}y(t)$ for $t \in [0, T]$, and $y(\cdot)$ and $u(\cdot)$ are the state and control variables valued in M and U , respectively. The cost functional associated with (1.1) is

$$J(u(\cdot)) = \int_0^T f^0(t, y(t), u(t)) dt. \quad (1.2)$$

In (1.2), $u(\cdot)$ belongs to the following admissible control set

$$\mathcal{U}_{ad} \equiv \{u(\cdot) : [0, T] \rightarrow U; \ u(\cdot) \text{ is measurable}\}. \quad (1.3)$$

Clearly, this is the situation without endpoint constraints.

We shall also consider the control system (1.1) with the following endpoint constraint

$$y(T) = y_1, \quad (1.4)$$

for some given $y_1 \in M$. The corresponding admissible control set is then given by

$$\mathcal{V}_{ad} \equiv \{u(\cdot) \in \mathcal{U}_{ad}; \ y_u(T) = y_1\},$$

where $y_u(\cdot)$ is the solution to (1.1) associated to the control $u(\cdot) \in \mathcal{U}_{ad}$.

In this work, we shall consider the following two optimal control problems:

Problem I To find a $\bar{u}(\cdot) \in \mathcal{U}_{ad}$ such that

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_{ad}} J(u(\cdot)); \quad (1.5)$$

Problem II To find a $\bar{u}(\cdot) \in \mathcal{V}_{ad}$ such that

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{V}_{ad}} J(u(\cdot)). \quad (1.6)$$

For each of the above problems, we call $\bar{u}(\cdot)$ an optimal control, the corresponding solution $\bar{y}(\cdot)$ to (1.1) an optimal trajectory, and $(\bar{y}(\cdot), \bar{u}(\cdot))$ an optimal pair. Clearly, each of the above two optimal control problems can be viewed as an optimal control problem with state constrained on a submanifold of the Euclidean space.

One of the central topics in control theory is to establish necessary and sufficient conditions for optimal controls. As that in calculus, one can derive the first-order necessary condition for optimal controls, as done in the classical monograph [32], even for some situation of state constraints. Nevertheless, for some optimal control problems, it may well happen that the first-order necessary conditions turn out to be trivial. In this case, the first-order necessary condition cannot provide enough information for the theoretical analysis and numerical computation, and therefore one needs to study the second (or even higher) order optimality conditions for optimal controls.

There are many works addressing the second order necessary and sufficient conditions for optimal control problems in Euclidean spaces, such as [2, 4, 8, 13, 14, 15, 16, 17, 20, 21, 23, 26, 27, 28, 29, 35] (in which [2, 8, 15, 21, 29] are five research monographs) and the references therein. Among them, we mention several works, in which the controls are subjected to some restrictions: Warga [35] considers a control system with the control set in a compact metric space. Frankowska and Tonon [13] investigate the second order necessary conditions when a smooth endpoint constraint is presented and the control set is a closed subset of an Euclidean space. For the case that the control set is a general metric space, Lou [26] considers a control system without state constraints. For the case that the state (or mixed state-control) satisfies inequality or equality constraints, we refer to [3, 19, 30] and so on.

For the control system whose state is constrained to a manifold, there are also some literatures (e.g. [1, 5, 6, 7, 33] and the references therein) devoted to the second order necessary and sufficient conditions for optimal controls when the control set is an open subset of some manifold. Agrachev and Sachkov show a Legendre-type second order necessary condition for optimal controls, i.e., the Hessian of the corresponding Hamiltonian function with respect to the control variable is semi-negative definite (see [1, Theorem 20.6, p. 300 and Proposition 20.11, p. 310]); also, they give a strong Legendre-type second order sufficient condition for an optimal controls for sufficiently short time intervals, i.e. the Hessian of the corresponding Hamiltonian function with respect to the control variable is negative definite (see [1, Proposition 20.12, p. 311 and Theorem 21.8, p. 347]). On the basis of the second order optimality conditions, Bonnard, Caillau and Trélat [5] present an algorithm to compute the first conjugate time along a smooth extremal curve, at which the trajectory fails to be optimal. In [33, Chapter 4], Schättler and Ledzewicz provide a “state of art” account on the second order necessary and sufficient optimality conditions for a class of control-affine systems (Also, we refer to [6, 7] for relatively earlier

works). Nevertheless, compared to the case of flat spaces, the picture of second optimality conditions in the curved spaces is, in our opinion, quite incomplete. For example, it would be quite interesting to extend the related results in the flat spaces (say, that in the research monographs [2, 8, 15, 21, 29] on second order optimality conditions) to the setting of curved spaces but many things remain to be done.

The main purpose of this paper is to investigate the second order necessary and sufficient conditions for optimal control problems evolved on Riemannian manifolds. We distinguish the problems into two cases. The first case is when the endpoint is free, i.e., the case without the condition (1.4). For this case, we consider the control systems with considerably general control sets. By means of the needle variation technique, we obtain the second order Taylor's expansion of the cost functional with respect to the control variable via the Riemannian geometric tools, introduce a second order dual equation which depends on the curvature tensor of the manifold, and obtain second order necessary conditions for optimal controls of integral and pointwise form respectively. In addition, we obtain the second order sufficient condition of integral form for local optimal controls. The second case is when the endpoint is fixed, i.e., the condition (1.4) is imposed. For this more difficult case, as in [1], we assume that the control set is open, and consider an extended control system involving an extra unknown variable associated to the cost functional. There is a well known relation between the optimal controls of the original optimal control problem and the above extended control system (e.g. [1, p. 179, Section 12.4]): the terminal value of the trajectory of the extended control system corresponding to an optimal control must be at the boundary of the attainable set of this control system at the terminal time. Thanks to this relation, Agrachev and Sachkov [1] observe that, the key point of finding second order optimality conditions for the original optimal controls is to analyze the Hessian of the endpoint mapping of the extended control system with respect to the control variable. Following this idea and employing again some tools from the global geometry, we compute an explicit expression of this Hessian (see (6.52)), and obtain the desired second order optimality conditions, which, as the case of free endpoint, contain also the curvature tensor along the optimal trajectory. Note that the above mentioned explicit expression is absent in [1].

From the viewpoint of Riemannian geometry, it is quite natural that the curvature tensor appears in the second order optimality condition, see Synge's second variation formula for geodesics (e.g., [31, Theorem 21, p. 158]). Nevertheless, one of the main difficulties for our control problems is to compute the second order Taylor's expansions (of the perturbed trajectories), i.e., (5.3)–(5.4) and (5.23), with convenient and explicit expressions of the second order terms, which, in principle, should involve the curvature tensor. To do this, quite different from the flat spaces, we need to establish a key geometric lemma, i.e, Lemma 2.3. This lemma enables us to obtain the second order variational

equations, i.e., (5.2) and (5.22), whose solutions are exactly the desired expressions of the above mentioned second order terms. Clearly, the curvature tensor does appear in these two equations. It seems to us that Lemma 2.3 has some independent interest and may be applied in other places. Another difficulty in our **Problem I** is how to introduce suitable dual equations such that the second order Taylor's expansion (6.5) (of the cost functional) can be rewritten in terms of the dual variables. This is the key to obtain the corresponding second order optimality condition. It is not difficult to single out the first order dual equation (3.7) (because, at least formally, it is the same as that for the flat spaces), which is employed to rewrite the first order term in (6.5) (see (6.6)). For the second order term (6.7), there are some quadratic terms of the first order variation $X^\epsilon(\cdot)$ (see (6.10)). Similar to the setting of flat spaces ([26]), one needs to find some second order dual equation to “cancel” these quadratic terms. Different from the flat setting and interestingly, the curvature tensor appears in our second order dual equation (3.8).

It is worth mentioning that, compared to the corresponding results in the Euclidean spaces ([2, 4, 8, 13, 14, 16, 17, 20, 21, 23, 26, 27, 28, 35]), the curvature tensor along the optimal trajectory appears precisely in our results (Theorems 3.1–3.4). This shows the very difference between the curved and the flat spaces from the viewpoint of optimal control theory. Also, our results for **Problem II** (i.e., Theorems 3.3–3.4) are different from [1, Theorem 20.6, p. 300, Proposition 20.11, p. 310 and Theorem 20.16, p. 317] and [1, Proposition 20.12, p. 311 and Theorem 21.8, p. 347]) at least in two aspects. Firstly, our second order optimality conditions depend on the curvature tensor, which does not appear explicitly in [1]. Secondly, our Theorem 3.4 guarantees the local optimality of a control in any time interval, while the corresponding result in [1] works only for a sufficiently short time interval.

In order to show the differences between our main results and that in the previous works, we shall provide two illustrative examples, i.e. Examples 4.1–4.2. Example 4.1 is about an optimal control problem on a non-compact manifold with negative curvature and with a discrete control set. We show that this problem fulfils all the assumptions in our main results. In Example 4.2, we apply one of our main results, i.e., Theorem 3.3, to the following famous geometric problem: Given any two points on a Riemannian manifold, find the necessary conditions of a locally shortest curve connecting these two points. We conclude that this curve must be a geodesic and satisfies the second variation of energy, which is consistent with the corresponding results in Riemannian geometry (see [31, p. 159]). However, the Legendre-type second order optimality condition in [1, Theorem 20.6, p. 300 and Proposition 20.11, p. 310] becomes trivial for the same problem (see Remark 4.1 for more detailed analysis).

Though geometric control theory is a huge subject and there exist many works on control theory via differential geometry, it seems to us that the more delicate tool, i.e.

Riemannian geometry, has rarely been used to the study of ordinary differential equation control problems. As far as we know, this work is the first one which links explicitly the curvature tensor of the manifold to the optimality conditions. It seems us that this may stimulate further work employing Riemannian geometry to the study of control problems.

The rest of this paper is organized as follows. In Section 2, for the reader's convenience we recall first some basic notions and results on Riemannian manifolds, and also we show two lemmas which will be useful later. The main results of this paper are stated in Section 3. In Section 4, we give the above mentioned two examples. In Section 5, we obtain the first and second order variations of a trajectory of the system (1.1), by means of two class of variation techniques, i.e., the needle variation and the classical variation. Sections 6 is devoted to proving our main results in this work.

One of the main results in this paper, i.e., Theorem 3.1, has been announced in [9] without proof.

2 Some preliminaries

2.1 Exponential map

For this part, we refer the readers to [12, Chapter 3] and [36, Chapter 3].

A differentiable curve $\gamma(t)$ on M with $t \in [0, \alpha)$ (for some $\alpha > 0$) is called a geodesic if it satisfies

$$\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t) = 0, \quad t \in [0, \alpha). \quad (2.1)$$

Let $x \in M$ be fixed. For any $v \in T_x M$, there exists a unique geodesic $\gamma_v(\cdot)$ satisfying that $\gamma_v(0) = x$ and $\dot{\gamma}_v(0) = v$. Let $[0, \ell_v)$ be the maximal interval on which $\gamma_v(\cdot)$ is defined. Let $O_x \subset T_x M$ be the set of vectors v such that $\ell_v > 1$. Then one can define the exponential map as follows

$$\exp_x : O_x \rightarrow M, \quad \exp_x v = \gamma_v(1).$$

It can be shown that O_x is a neighborhood of the origin $O \in T_x M$, and \exp_x maps straight line segments in $T_x M$ passing through the origin $O \in T_x M$ to geodesic segments in M passing through x . For any $v \in T_x M$, the differential of \exp_x at v is a linear map, denoted by

$$d\exp_x|_v : T_v T_x M \rightarrow T_{\exp_x v} M, \quad (2.2)$$

where $T_v T_x M$ denotes the tangent space of the manifold $T_x M$ at the point $v \in T_x M$.

Given an $\epsilon > 0$, write

$$B(O, \epsilon) \equiv \{v \in T_x M; \quad |v| < \epsilon\} \quad \text{and} \quad B_x(\epsilon) \equiv \{y \in M; \quad \rho(x, y) < \epsilon\}. \quad (2.3)$$

We call $i(x) \equiv \sup\{\epsilon > 0; \quad \text{The map } \exp_x : B(O, \epsilon) \rightarrow B_x(\epsilon) \text{ is diffeomorphic}\}$ the injectivity radius at the point x (e.g., [31, p. 142]).

We list the following properties of the exponential map, which can be found in many books on Riemannian geometry (e.g. the proof of [12, Proposition 2.9, p. 65]).

Lemma 2.1 *For any $x \in M$, the map \exp_x is a local diffeomorphism, whose differential at the origin $O \in T_x M$ satisfies*

$$d\exp_x|_O = d\exp_x^{-1}|_x = \text{the identity operator on } T_x M. \quad (2.4)$$

Furthermore, for any $y \in M$ with $\rho(x, y) < i(x)$, there exists a unique shortest piecewise smooth curve which is also a geodesic in M , connecting x and y .

2.2 Parallel translation and tensors

For the details of this part, we refer the readers to [22, Chapter I and Chapter III], [31, Chapter 2], [36, Chapter 1] and [18, Chapter 1].

For any $x \in M$ and $r, s \in \mathbb{N}$, a multilinear map

$$F : \underbrace{T_x^* M \times \cdots \times T_x^* M}_r \times \underbrace{T_x M \times \cdots \times T_x M}_s \rightarrow \mathbb{R}$$

is called a tensor of order (r, s) at x . Denote by $\mathcal{T}_s^r(x)$ the tensor space of type (r, s) at x . A tensor field \mathcal{T} of type (r, s) on M is a smooth assignment of a tensor $\mathcal{T}(x) \in \mathcal{T}_s^r(x)$ to each point x of M . The norm of \mathcal{T} at $x \in M$ is defined as follows:

$$\begin{aligned} |\mathcal{T}(x)| &= \sup \{ \mathcal{T}(x)(Y_1, \dots, Y_r, \lambda_1, \dots, \lambda_s); \ Y_j \in T_x^* M, \lambda_l \in T_x M, \\ &|Y_j| \leq 1, |\lambda_l| \leq 1, j = 1, \dots, r, l = 1, \dots, s \}, \quad x \in M. \end{aligned} \quad (2.5)$$

Denote by $\mathcal{T}_s^r(M)$ the set of all tensor fields of type (r, s) over M .

Let $\gamma : [0, \ell] \rightarrow M$ be a differentiable curve with $\gamma(0) = x \in M$, $\gamma(\ell) = y \in M$ and $\ell > 0$. Given a vector $v \in T_x M$, there exists a unique vector field X along γ satisfying

$$\nabla_{\dot{\gamma}(s)} X = 0, \quad \forall s \in [0, \ell], \quad X(0) = v.$$

The mapping $T_x M \ni v \mapsto X(\gamma(\ell)) \in T_y M$ is a linear isometry between $T_x M$ and $T_y M$. We call this map the parallel translation along the curve γ , and denote it by $L_{xy}^\gamma v$. The parallel translation along the curve γ enjoys the following property:

$$\langle L_{xy}^\gamma X_1, L_{xy}^\gamma X_2 \rangle = \langle X_1, X_2 \rangle, \quad \forall X_1, X_2 \in T_x M. \quad (2.6)$$

For any $\eta \in T_x^* M$, we define $L_{xy}^\gamma \eta \in T_y^* M$ by $L_{xy}^\gamma \eta(X) = \eta((L_{xy}^\gamma)^{-1} X)$ for any $X \in T_y M$. One can extend the parallel translation of a vector at $x \in M$ along the curve γ to a tensor $\mathcal{T} \in \mathcal{T}_s^r(x)$ by

$$L_{xy}^\gamma \mathcal{T}(v_1, \dots, v_r, \eta_1, \dots, \eta_s) = \mathcal{T}((L_{xy}^\gamma)^{-1} v_1, \dots, (L_{xy}^\gamma)^{-1} v_r, (L_{xy}^\gamma)^{-1} \eta_1, \dots, (L_{xy}^\gamma)^{-1} \eta_s),$$

for all $v_1, \dots, v_r \in T_y^*M$ and $\eta_1, \dots, \eta_s \in T_yM$. From the above formula, one can get

$$\mathcal{T}(v_1, \dots, v_r, \eta_1, \dots, \eta_s) = L_{xy}^\gamma \mathcal{T}(L_{xy}^\gamma v_1, \dots, L_{xy}^\gamma v_r, L_{xy}^\gamma \eta_1, \dots, L_{xy}^\gamma \eta_s), \quad (2.7)$$

for any $v_1, \dots, v_r \in T_x^*M$ and $\eta_1, \dots, \eta_s \in T_xM$. Especially, from (2.5), (2.6) and (2.7), it follows that

$$|X| = |L_{xy}^\gamma X|, \quad |\lambda| = |L_{xy}^\gamma \lambda|, \quad \forall X \in T_xM, \quad \lambda \in T_x^*M. \quad (2.8)$$

In particular, if $\rho(x, y) < \min\{i(x), i(y)\}$, according to Lemma 2.1, there is a unique shortest geodesic γ connecting x and y . In this case, we use L_{xy} instead of L_{xy}^γ for abbreviation.

Let \mathcal{T} be a tensor field. Take any $v \in T_xM$. Let γ be a smooth curve such that $\gamma(0) = x$ and $\dot{\gamma}(0) = v$. Then the covariant derivative of a tensor field (in terms of parallel translation) is defined as follows (see [18, p. 42]):

$$\nabla_v \mathcal{T} = \lim_{t \rightarrow 0} \frac{1}{t} \left(L_{\gamma(t)x}^\gamma \mathcal{T}(\gamma(t)) - \mathcal{T}(x) \right). \quad (2.9)$$

Denote by $\nabla \mathcal{T}$ the covariant differential of \mathcal{T} , which is a tensor field of order $(r, s+1)$, and is defined by (see [22, p. 124])

$$\nabla \mathcal{T}(Y_1, \dots, Y_r, \lambda_1, \dots, \lambda_s, Z) = \nabla_Z \mathcal{T}(Y_1, \dots, Y_r, \lambda_1, \dots, \lambda_s), \quad (2.10)$$

for all $Y_1, \dots, Y_r \in T^*M$ and $\lambda_1, \dots, \lambda_s, Z \in TM$. Applying [22, Proposition 2.7, p. 123 and Proposition 2.10, p. 124], one can get the following property: For each $Z \in TM$, one has

$$\begin{aligned} & \nabla_Z \mathcal{T}(Y_1, \dots, Y_r, \lambda_1, \dots, \lambda_s) \\ &= Z \left(\mathcal{T}(Y_1, \dots, Y_r, \lambda_1, \dots, \lambda_s) \right) - \mathcal{T}(\nabla_Z Y_1, Y_2, \dots, Y_r, \lambda_1, \dots, \lambda_s) \\ & \quad - \dots - \mathcal{T}(Y_1, \dots, \nabla_Z Y_r, \lambda_1, \dots, \lambda_s) - \mathcal{T}(Y_1, \dots, Y_r, \nabla_Z \lambda_1, \lambda_2, \dots, \lambda_s) - \dots \\ & \quad - \mathcal{T}(Y_1, \dots, Y_r, \lambda_1, \dots, \nabla_Z \lambda_s), \quad \forall Y_1, \dots, Y_r \in T^*M, \quad \lambda_1, \dots, \lambda_s \in TM. \end{aligned} \quad (2.11)$$

In particular, a smooth function $f \in C^\infty(M)$ is a tensor of order $(0, 0)$. $\nabla^2 f$ is a tensor of order $(0, 2)$. We call this tensor the Hessian of the function f , which is a symmetric tensor, and can be computed by

$$\nabla^2 f(x)(X, Y) = Y(x)(Xf) - (\nabla_{Y(x)} X)f, \quad x \in M, \quad X, Y \in TM. \quad (2.12)$$

For a smooth function $h : M \times M \rightarrow \mathbb{R}$ of two arguments, we denote by $\nabla_i h$ the covariant derivative of h with respect to the i^{th} argument with $i = 1, 2$, i.e., for $X \in TM$ and $(x_1, x_2) \in M \times M$,

$$\langle \nabla_i h(x_1, x_2), X(x_i) \rangle = X(x_i)h(x_1, x_2). \quad (2.13)$$

Moreover, we define higher order derivatives of h as follows: For $i, j = 1, 2, i \neq j$, any $(x_1, x_2) \in M \times M$ and $X, Y, Z \in TM$,

$$\begin{aligned}\nabla_i \nabla_j h(x_1, x_2)(X, Y) &\equiv Y(x_i) \left(X(x_j) (h(x_1, x_2)) \right) = Y(x_i) (\langle \nabla_j h(x_1, x_2), X(x_j) \rangle); \\ \nabla_i^2 h(x_1, x_2)(X, Y) &\equiv Y(x_i) \left(X(x_i) h(x_1, x_2) \right) - \nabla_{Y(x_i)} X h(x_1, x_2); \\ \nabla_i^2 \nabla_j h(x_1, x_2)(X, Y, Z) &\equiv \nabla_i^2 (\langle X(x_j), \nabla_j h(x_1, x_2) \rangle)(Y, Z); \\ \nabla_i \nabla_j^2 h(x_1, x_2)(X, Y, Z) &\equiv Z(x_i) \left(\nabla_j^2 h(x_1, x_2)(X, Y) \right).\end{aligned}\tag{2.14}$$

Let us recall the definition of the product of tensors (see [22, p.22]):

$$\mathcal{T} \otimes \mathcal{K} = X_1 \otimes \cdots \otimes X_r \otimes Y_1 \otimes \cdots \otimes Y_p \otimes \omega_1 \otimes \cdots \otimes \omega_s \otimes \eta_1 \otimes \cdots \otimes \eta_q \in \mathcal{T}_{s+q}^{r+p}(M), \tag{2.15}$$

where $\mathcal{T} = X_1 \otimes \cdots \otimes X_r \otimes \omega_1 \otimes \cdots \otimes \omega_s \in \mathcal{T}_s^r(M)$, $\mathcal{K} = Y_1 \otimes \cdots \otimes Y_p \otimes \eta_1 \otimes \cdots \otimes \eta_q \in \mathcal{T}_q^p(M)$ and $r, s, p, q \in \mathbb{N}$. For any $i, j \in \mathbb{N}$, denote by \mathcal{E}^{ij} the contraction of the i th contravariant index and the j th covariant index, which is a linear mapping from $\mathcal{T}_s^r(M)$ to $\mathcal{T}_{s-1}^{r-1}(M)$ with $1 \leq i \leq r$ and $1 \leq j \leq s$, and is defined by (see [18, p.17])

$$\begin{aligned}\mathcal{E}^{ij}(X_1 \otimes \cdots \otimes X_r \otimes \omega_1 \otimes \cdots \otimes \omega_s) \\ = \omega_j(X_i) X_1 \otimes \cdots \otimes X_{i-1} \otimes X_{i+1} \otimes \cdots \otimes X_r \otimes \omega_1 \otimes \cdots \otimes \omega_{j-1} \otimes \omega_{j+1} \otimes \cdots \otimes \omega_s,\end{aligned}\tag{2.16}$$

for all $X_1, \cdots, X_r \in TM$ and $\omega_1, \cdots, \omega_s \in T^*M$.

2.3 Two useful lemmas

We begin with the following technical result (in which, (2.18)–(2.19) and the first equality in (2.17) can be found in [10]).

Lemma 2.2 *For any $x, y \in M$ with $\rho(x, y) < \min\{i(x), i(y)\}$, $X, X_1, X_2 \in T_x M$ and $Y \in T_y M$, it holds that*

$$|\exp_x^{-1} y| = |\exp_y^{-1} x| = \rho(x, y), \quad \nabla_{X_1} L_x X = 0, \tag{2.17}$$

$$\nabla_1 \rho^2(x, y) = -2 \exp_x^{-1} y, \quad \nabla_2 \rho^2(x, y) = -2 \exp_y^{-1} x, \tag{2.18}$$

$$L_{xy} \exp_x^{-1} y = -\exp_y^{-1} x, \quad L_{xy} d_1 \rho^2(x, y) = -d_1 \rho^2(y, x), \tag{2.19}$$

$$\nabla_1 \nabla_2 \rho^2(x, y)(Y, X) = -2 \langle d \exp_y^{-1} |_x X, Y \rangle, \tag{2.20}$$

$$\langle d \exp_x^{-1} |_y Y, X \rangle = \langle d \exp_y^{-1} |_x X, Y \rangle, \tag{2.21}$$

$$\nabla_1 \nabla_2 \rho^2(x, y)(Y, X) = -\nabla_1^2 \rho^2(x, y)(L_{yx} Y, X) - \langle \nabla_1 \rho^2(x, y), \nabla_X L_y Y \rangle, \tag{2.22}$$

$$\nabla_1^2 \rho^2(x, x)(X_1, X_2) = \nabla_2^2 \rho^2(x, x)(X_1, X_2) = 2 \langle X_1, X_2 \rangle, \tag{2.23}$$

$$\nabla_i \nabla_j^2 \rho^2(x, x) = \nabla_i^2 \nabla_j \rho^2(x, x) = \nabla_i^3 \rho^2(x, x) = 0, \quad i, j = 1, 2, i \neq j, \tag{2.24}$$

where the notions $\nabla_1 \nabla_2 \rho^2$, $\nabla_i^2 \nabla_j \rho^2$ and $\nabla_i \nabla_j^2 \rho^2$ with $i, j = 1, 2$ and $i \neq j$ are defined in (2.14), $\nabla_i^2 \rho^2$ is the Hessian of ρ^2 with respect to the i^{th} argument, $\nabla_i^3 \rho^2$ is the covariant derivative of the Hessian $\nabla_i^2 \rho^2(x, x)$ with respect to the i^{th} argument (see (2.10)), and d_i stands for the exterior derivative of a function on $M \times M$ with respect to the i^{th} argument for $i = 1, 2$.

Proof: We only prove the second equality in (2.17), and (2.20)–(2.24).

Let γ_1 be a radial geodesic satisfying $\gamma_1(0) = x$ and $\dot{\gamma}_1(0) = X_1$. Then, γ_1 is the shortest geodesic connecting $\gamma_1(s)$ and x , provided that $s > 0$ is small enough. By (2.9), we have

$$\nabla_{X_1} L_{x \cdot} X = \lim_{s \rightarrow 0^+} \frac{1}{s} \left(L_{\gamma_1(s)x} L_{x\gamma_1(s)} X - X \right) = 0,$$

which gives the second equality of (2.17).

Let γ and β be the curves satisfying

$$\gamma(0) = x, \quad \dot{\gamma}(0) = X, \quad \beta(0) = y, \quad \dot{\beta}(0) = Y. \quad (2.25)$$

By (2.14) and noting (2.18), we have

$$\begin{aligned} \nabla_1 \nabla_2 \rho^2(x, y)(Y, X) &= \frac{\partial}{\partial \theta} \frac{\partial}{\partial \tau} \rho^2(\gamma(\theta), \beta(\tau)) \Big|_{\theta=\tau=0} = \frac{\partial}{\partial \theta} \langle \nabla_2 \rho^2(\gamma(\theta), \beta(0)), \dot{\beta}(0) \rangle \Big|_{\theta=0} \\ &= -2 \frac{\partial}{\partial \theta} \langle \exp_y^{-1} \gamma(\theta), Y \rangle \Big|_{\theta=0} = -2 \langle d \exp_y^{-1} |_x X, Y \rangle, \end{aligned}$$

which implies (2.20). Based on the above identity, (2.18), (2.19) and (2.6), we have

$$\begin{aligned} \nabla_1 \nabla_2 \rho^2(x, y)(Y, X) &= -2 \frac{\partial}{\partial \theta} \langle L_{y\gamma(\theta)} \exp_y^{-1} \gamma(\theta), L_{y\gamma(\theta)} Y \rangle \Big|_{\theta=0} \\ &= -\frac{\partial}{\partial \theta} \langle \nabla_1 \rho^2(\gamma(\theta), y), L_{y\gamma(\theta)} Y \rangle \Big|_{\theta=0} \\ &= -\langle \nabla_X \nabla_1 \rho^2(\cdot, y), L_{yx} Y \rangle - \langle \nabla_1 \rho^2(x, y), \nabla_X L_y \cdot Y \rangle, \end{aligned}$$

which, together with (2.10), implies (2.22).

To prove (2.21), we recall (2.25) and the definition of the differential of a differentiable map, apply (2.18) and get

$$\begin{aligned} \langle d \exp_x^{-1} |_y Y, X \rangle &= \langle d \exp_x^{-1} |_y \dot{\beta}(0), X \rangle = \left\langle \frac{d}{d\tau} \exp_x^{-1} \beta(\tau) \Big|_{\tau=0}, X \right\rangle \\ &= -\frac{1}{2} \frac{d}{d\tau} \langle \nabla_1 \rho^2(x, \beta(\tau)), \dot{\gamma}(0) \rangle \Big|_{\tau=0} \\ &= -\frac{1}{2} \frac{d}{d\tau} \dot{\gamma}(0) \left(\rho^2(\cdot, \beta(\tau)) \right) \Big|_{\tau=0} = -\frac{1}{2} \frac{\partial}{\partial \tau} \frac{\partial}{\partial \theta} \rho^2(\gamma(\theta), \beta(\tau)) \Big|_{\theta=\tau=0}. \end{aligned}$$

Similarly, we get $\langle d \exp_y^{-1} |_x X, Y \rangle = -\frac{1}{2} \frac{\partial}{\partial \tau} \frac{\partial}{\partial \theta} \rho^2(\gamma(\theta), \beta(\tau)) \Big|_{\theta=\tau=0}$. These two equalities imply (2.21).

To prove (2.23), we choose the normal coordinates $\{x_1, \dots, x_n\}$ at x such that

$$\left\langle \frac{\partial}{\partial x_i}(x), \frac{\partial}{\partial x_j}(x) \right\rangle = \delta_{ij}, \quad \nabla_{\frac{\partial}{\partial x_i}(x)} \frac{\partial}{\partial x_j} = 0, \quad (2.26)$$

for $i, j = 1, \dots, n$. From the above property, we can further deduce that

$$\frac{\partial}{\partial x_i} \Big|_x g_{kl}(\cdot) = 0, \quad (2.27)$$

with $g_{kl}(\cdot) \equiv \langle \frac{\partial}{\partial x_k}(\cdot), \frac{\partial}{\partial x_l}(\cdot) \rangle$, where $i, k, l = 1, \dots, n$. By some computation in this system of local coordinates, we obtain (2.23).

Then we go to the proof of (2.24). For any y which is sufficiently close to x and any vector field F , recalling (2.14), we have

$$\nabla_1^2 \nabla_2 \rho^2(y, x)(F, \cdot, \cdot) = \nabla_1^2 \left(\nabla_2 \rho^2(y, x)(F) \right)(\cdot, \cdot) = -\nabla_1^2 \left(\nabla_1 \rho^2(y, x)(L_{xy}F) \right)(\cdot, \cdot),$$

where we have used (2.7), (2.18) and (2.19). Let us compute the right hand side of the above identity in local coordinates. By the property of the normal coordinates $\{x_1, \dots, x_n\}$, we have $\rho^2(x, y) = |\exp_x^{-1} y|^2 = \sum_{i=1}^n x_i^2(y)$. For any z which is sufficiently close to x , denote

$$L_{zy}F = \sum_{i=1}^n a_i(z, y) \frac{\partial}{\partial x_i}(y), \quad (2.28)$$

where $a_i(z, y)$ is a function depending on z and y . Then,

$$\nabla_1 \rho^2(y, x)(L_{xy}F) = L_{xy}F \left(\rho^2(\cdot, x) \right) = \sum_{i,j=1}^n a_j(x, y) \frac{\partial}{\partial x_j}(y) (x_i^2(\cdot)) = 2 \sum_{i=1}^n a_i(x, y) x_i(y).$$

For any vector fields $X_j(\cdot) = \sum_{k=1}^n X_j^k(\cdot) \frac{\partial}{\partial x_k}(\cdot)$ with $j = 1, 2$, noting (2.12), (2.19) and (2.26), we obtain that

$$\nabla_1^2 \nabla_2 \rho^2(x, x)(F, X_1, X_2) = -2 \sum_{k,l,m=1}^n X_2^k(x) X_1^l(x) \frac{\partial^2}{\partial x_k \partial x_l} \Big|_x \left(a_m(x, \cdot) x_m(\cdot) \right).$$

We claim that

$$\sum_{k,l,m=1}^n X_2^k(x) X_1^l(x) \frac{\partial^2}{\partial x_k \partial x_l} \Big|_x \left(a_m(x, \cdot) x_m(\cdot) \right) = 0. \quad (2.29)$$

In fact, according to (2.26), (2.27) and (2.28) we have

$$\begin{aligned} \frac{\partial}{\partial x_k} \Big|_x a_m(x, \cdot) &= \sum_{i=1}^n \frac{\partial}{\partial x_k} \Big|_x \left(a_i(x, \cdot) g_{im}(\cdot) \right) = \frac{\partial}{\partial x_k} \Big|_x \left\langle L_x F, \frac{\partial}{\partial x_m}(\cdot) \right\rangle \\ &= \left\langle \nabla_{\frac{\partial}{\partial x_k} \Big|_x} L_x F, \frac{\partial}{\partial x_m} \Big|_x \right\rangle, \end{aligned}$$

where $k = 1, \dots, n$. By the second equality of (2.17), we have $\frac{\partial}{\partial x_k} \Big|_x a_m(x, \cdot) = 0$, which implies $\nabla_1^2 \nabla_2 \rho^2(x, x) = 0$. Using the same method, we can show that $\nabla_2 \nabla_1^2 \rho^2(x, x) = 0$,

$\nabla_1 \nabla_2^2 \rho^2(x, x) = 0$ and $\nabla_2^2 \nabla_1 \rho^2(x, x) = 0$. We now prove that $\nabla_1^3 \rho^2(x, x) = 0$. For this purpose, we take any $y \in M$ which is sufficiently close to x . For any $Y \in T_y M$ and $j, k = 1, \dots, n$, letting $\frac{\partial}{\partial x_k}(x)$ act on both sides of (2.22) with $X = \frac{\partial}{\partial x_j}$, we obtain that

$$\begin{aligned} & \nabla_1^2 \nabla_2 \rho^2(x, y) \left(Y, \frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k} \right) \\ &= -\nabla_1^3 \rho^2(x, y) \left(L_{y*} Y, \frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k} \right) - \nabla_1^2 \rho^2(x, y) \left(\nabla_{\frac{\partial}{\partial x_k}} L_{y*} Y, \frac{\partial}{\partial x_j} \right) \\ & \quad - \nabla_1^2 \rho^2(x, y) \left(\nabla_{\frac{\partial}{\partial x_j}} L_{y*} Y, \frac{\partial}{\partial x_k} \right) - \langle \nabla_1 \rho^2(x, y), \nabla_{\frac{\partial}{\partial x_k}} \nabla_{\frac{\partial}{\partial x_j}} L_{y*} Y \rangle, \end{aligned}$$

where we have used (2.11) and (2.26). Letting y approach to x , and using (2.17), (2.18) and $\nabla_1^2 \nabla_2 \rho^2(x, x) = 0$, we conclude that $\nabla_1^3 \rho^2(x, x) = 0$. Similarly, we can get $\nabla_2^3 \rho^2(x, x) = 0$ in the same way. \square

Denote by $[X, Y] \equiv XY - YX$ the Lie bracket of vector fields X and Y . Denote by R the curvature tensor (of (M, g)), which is a correspondence that associates to every pair $X, Y \in TM$ a mapping $R(X, Y) : TM \rightarrow TM$ given by

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z, \quad \forall Z \in TM.$$

We write

$$R(X, Y, Z, W) = \langle R(X, Y)Z, W \rangle, \quad \forall X, Y, Z, W \in TM.$$

For any $x \in M$ and $X(x), Y(x) \in T_x M$ with $X(x) \nparallel Y(x)$, the sectional curvature $\text{sec}(X(x), Y(x))$ (of the plane spanned by $X(x)$ and $Y(x)$) is given by

$$\text{sec}(X(x), Y(x)) \equiv \frac{R(X, Y, Y, X)(x)}{|X(x) \wedge Y(x)|^2}. \quad (2.30)$$

The following result will play a key role in the sequel.

Lemma 2.3 *Let $X, V, F \in TM$. Then, for any $x \in M$, it holds that*

$$\nabla_{X(x)} \left(d_x(L_{x*} F(x)) V(x) \right) = \frac{1}{2} R(X, V) F(x)^1; \quad (2.31)$$

$$\nabla_2^3 \nabla_1 \rho^2(x, x)(X, F, V, V) + \nabla_2^2 \nabla_1^2 \rho^2(x, x)(X, F, V, V) = 2R(X, V, F, V)(x). \quad (2.32)$$

¹The left hand side of (2.31) is defined as follows: For $y \in M$ closing enough to x , we view $L_{y*} F(\cdot) : B_x(i(x)) \rightarrow T_y M$ as a map, where $B_x(i(x))$ is defined in (2.3). The differential of $L_{y*} F(\cdot)$ at x is a linear map, denoted by $d_x(L_{xy} F(x)) : T_x M \rightarrow T_{L_{xy} F(x)} T_y M$, where $T_{L_{xy} F(x)} T_y M$ is isomorphic to $T_y M$. Therefore, we have $d_x(L_{xy} F(x)) V(x) \in T_y M$. Letting y vary around x , we get a vector field $d_x(L_{x*} F(x)) V(x)$ around x . Hence, $\nabla_{X(x)} \left(d_x(L_{x*} F(x)) V(x) \right)$ is the covariant derivative of vector field $d_x(L_{x*} F(x)) V(x)$ relative to $X(x)$.

Proof: We will prove (2.31) and (2.32) by means of suitable local coordinates. Let $\{x_1, \dots, x_n\}$ be the normal coordinates in a neighborhood $O \subset M$ around x , which satisfies (2.26) and (2.27). Denote by $(g_{ij})(\cdot) \equiv \left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle(\cdot)$ and Γ_{ij}^k ($i, j, k = 1, \dots, n$) the metric matrix and the Christoffel symbols, respectively. Take any $m, j \in \{1, \dots, n\}$. Let $x_j(\cdot)$ be the coordinate curve such that $\frac{\partial}{\partial x_j}(x) = \dot{x}_j(0)$ (See [12, p. 8]). Recalling (2.13), (2.28) and the definition of the differential of a differentiable map, for any $y \in O$, we have

$$d_x(L_{xy}F(x))\frac{\partial}{\partial x_j}(x) = \frac{\partial}{\partial s}\Big|_0 L_{x_j(s)y}F(x_j(s)) = \sum_{i=1}^n \langle \nabla_1 a_i(x, y), \frac{\partial}{\partial x_j}(x) \rangle \frac{\partial}{\partial x_i}(y),$$

where $\nabla_l a_i(\cdot, \cdot)$ (with $l = 1, 2$ and $i = 1, \dots, n$) is defined in (2.13). Then, recalling (2.26) and (2.14), we have

$$\nabla_{\frac{\partial}{\partial x_m}(x)} \left(d_x(L_{xy}F(x))\frac{\partial}{\partial x_j}(x) \right) = \sum_{i=1}^n \nabla_2 \nabla_1 a_i(x, x) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_m} \right) \frac{\partial}{\partial x_i}(x). \quad (2.33)$$

Clearly, to obtain (2.31), we need to compute $\nabla_2 \nabla_1 a_i(x, x) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_m} \right)$ for $i, j, m = 1, \dots, n$. To this end, we first show some properties of short geodesics. Given $z, y \in O$, write $V \equiv \exp_z^{-1} y$ and $\gamma_{zy}(s) \equiv \exp_z sV$ with $s \in [0, 1]$. Then, $\gamma_{zy}(\cdot)$ is the shortest geodesic connecting z and y , and satisfies $\gamma_{zy}(0) = z$, $\gamma_{zy}(1) = y$ and $\dot{\gamma}_{zy}(0) = \exp_z^{-1} y$. We assume that the local expression of $\gamma_{zy}(s)$ with $s \in [0, 1]$ in the system of coordinates (O, x_1, \dots, x_n) is as follows:

$$(\gamma_{zy}^1(s), \dots, \gamma_{zy}^n(s)).$$

Then, $\dot{\gamma}_{zy}(s) = \sum_{j=1}^n \dot{\gamma}_{zy}^j(s) \frac{\partial}{\partial x_j}(\gamma_{zy}(s))$. In particular, by Lemma 2.2, recalling the definition of a geodesic and using the local expression of $\gamma_{zy}(\cdot)$, we have

$$\dot{\gamma}_{zy}(1) = \sum_{i=1}^n \dot{\gamma}_{zy}^i(1) \frac{\partial}{\partial x_i}(y) = L_{zy} \dot{\gamma}_{zy}(0) = L_{zy} \exp_z^{-1} y = -\frac{1}{2} L_{zy} \nabla_1 \rho^2(z, y) = \frac{1}{2} \nabla_2 \rho^2(z, y), \quad (2.34)$$

which implies that $\dot{\gamma}_{zy}^i(1)$ with $i = 1, \dots, n$ are determined by the endpoints of geodesic $\gamma_{zy}(\cdot)$. Thus, we may view $\dot{\gamma}_{zy}^1(1), \dots, \dot{\gamma}_{zy}^n(1)$ as functions of the first argument z and the second argument y .

We claim that (Recall (2.13) and (2.14) for the notations)

$$\langle \nabla_2 \dot{\gamma}_{xy}^j(1), \frac{\partial}{\partial x_\zeta}(y) \rangle = \delta_j^\zeta + o(1), \quad \langle \nabla_1 \dot{\gamma}_{xy}^j(1), \frac{\partial}{\partial x_\zeta}(x) \rangle = -\delta_j^\zeta + o(1), \quad (2.35)$$

$$\nabla_1 \nabla_2 \dot{\gamma}_{xy}^j(1) = o(1), \quad \nabla_2^2 \dot{\gamma}_{xy}^j(1) = o(1), \quad (2.36)$$

for $\zeta, j = 1, \dots, n$, where δ_j^ζ is the usual Kronecker symbol, $o(1)$ is a tensor of suitable order satisfying $\lim_{y \rightarrow x} o(1) = 0$. Indeed, for any $j = 1, \dots, n$, it follows from (2.34) that

$$\frac{1}{2} \langle \nabla_2 \rho^2(z, y), \frac{\partial}{\partial x_j}(y) \rangle = \sum_{i=1}^n \langle \dot{\gamma}_{zy}^i(1), \frac{\partial}{\partial x_i}(y), \frac{\partial}{\partial x_j}(y) \rangle. \quad (2.37)$$

Letting $\frac{\partial}{\partial x_\zeta}(y)$ act on the above identity, by (2.11), we get

$$\begin{aligned} & \frac{1}{2} \nabla_2^2 \rho^2(z, y) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_\zeta} \right) + \frac{1}{2} \langle \nabla_2 \rho^2(z, y), \nabla_{\frac{\partial}{\partial x_\zeta}(y)} \frac{\partial}{\partial x_j} \rangle \\ &= \sum_{i=1}^n \left(\langle \nabla_2 \dot{\gamma}_{zy}^i(1), \frac{\partial}{\partial x_\zeta}(y) \rangle g_{ij}(y) + \dot{\gamma}_{zy}^i(1) \frac{\partial}{\partial x_\zeta}(y) g_{ij} \right). \end{aligned} \quad (2.38)$$

According to (2.17), (2.18), (2.23), (2.24), (2.26), (2.27) and the fact that

$$\dot{\gamma}_{xx}^i(1) = 0, \quad i = 1, \dots, n, \quad (2.39)$$

we obtain the first identity of (2.35) by letting $z = x$ in (2.38). In a similar way, we obtain the second identity of (2.35) by letting $\frac{\partial}{\partial x_\zeta}(z)$ act on (2.37) at $z = x$ with $\zeta = 1, \dots, n$. For $m = 1, \dots, n$, letting $\frac{\partial}{\partial x_m}(z)$ act on (2.38), via (2.11), we get

$$\begin{aligned} & \frac{1}{2} \nabla_1 \nabla_2^2 \rho^2(z, y) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_\zeta}, \frac{\partial}{\partial x_m}(z) \right) + \frac{1}{2} \nabla_1 \nabla_2 \rho^2(z, y) \left(\nabla_{\frac{\partial}{\partial x_\zeta}(y)} \frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_m}(z) \right) \\ &= \sum_{i=1}^n \left(\nabla_1 \nabla_2 \dot{\gamma}_{zy}^i(1) \left(\frac{\partial}{\partial x_\zeta}(y), \frac{\partial}{\partial x_m}(z) \right) g_{ij}(y) + \langle \nabla_1 \dot{\gamma}_{zy}^i(1), \frac{\partial}{\partial x_m}(z) \rangle \frac{\partial}{\partial x_\zeta}(y) g_{ij} \right). \end{aligned}$$

By Lemma 2.2, (2.26), (2.27) and (2.35), we obtain the first equality of (2.36) by taking $z = x$ in the above identity. Similarly, we obtain the second equality of (2.36) by letting $\frac{\partial}{\partial x_m}(y)$ ($m = 1, \dots, n$) act on (2.38) with $z = x$.

Recalling the definition of parallel translation and (2.28), for any $s \in [0, 1]$, we have

$$\begin{aligned} 0 &= \nabla_{\dot{\gamma}_{zy}(s)} L_z F(z) = \sum_{i, \zeta=1}^n \dot{\gamma}_{zy}^i(s) \nabla_{\frac{\partial}{\partial x_i}(\gamma_{zy}(s))} \left(a_\zeta(z, \cdot) \frac{\partial}{\partial x_\zeta} \right) \\ &= \sum_{\zeta=1}^n \left(\sum_{i=1}^n \dot{\gamma}_{zy}^i(s) \langle \nabla_2 a_\zeta(z, \gamma_{zy}(s)), \frac{\partial}{\partial x_i} \rangle \right. \\ &\quad \left. + \sum_{i, \eta=1}^n \dot{\gamma}_{zy}^i(s) a_\eta(z, \gamma_{zy}(s)) \Gamma_{i\eta}^\zeta(\gamma_{zy}(s)) \right) \frac{\partial}{\partial x_\zeta}(\gamma_{zy}(s)), \end{aligned}$$

which indicates that

$$\sum_{i=1}^n \dot{\gamma}_{zy}^i(1) \langle \nabla_2 a_\zeta(z, y), \frac{\partial}{\partial x_i} \rangle + \sum_{i, \eta=1}^n \dot{\gamma}_{zy}^i(1) a_\eta(z, y) \Gamma_{i\eta}^\zeta(y) = 0, \quad \zeta = 1, \dots, n.$$

Letting $\frac{\partial}{\partial x_m}(y)$ act on both sides of the above identity, via (2.11), we obtain that

$$\begin{aligned} & \sum_{i=1}^n \left(\langle \nabla_2 \dot{\gamma}_{zy}^i(1), \frac{\partial}{\partial x_m}(y) \rangle \langle \nabla_2 a_\zeta(z, y), \frac{\partial}{\partial x_i}(y) \rangle + \dot{\gamma}_{zy}^i(1) \nabla_2^2 a_\zeta(z, y) \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_m} \right) \right. \\ & \quad \left. + \dot{\gamma}_{zy}^i(1) \langle \nabla_2 a_\zeta(z, y), \nabla_{\frac{\partial}{\partial x_m}(y)} \frac{\partial}{\partial x_i} \rangle \right) + \sum_{i, \eta=1}^n \left(\langle \nabla_2 \dot{\gamma}_{zy}^i(1), \frac{\partial}{\partial x_m}(y) \rangle a_\eta(z, y) \Gamma_{i\eta}^\zeta(y) \right. \\ & \quad \left. + \dot{\gamma}_{zy}^i(1) \langle \nabla_2 a_\eta(z, y), \frac{\partial}{\partial x_m}(y) \rangle \Gamma_{i\eta}^\zeta(y) + \dot{\gamma}_{zy}^i(1) a_\eta(z, y) \frac{\partial}{\partial x_m}(y) \Gamma_{i\eta}^\zeta(y) \right) \\ &= 0. \end{aligned} \quad (2.40)$$

Letting $\frac{\partial}{\partial x_j}(z)$ act on the above identity at $z = x$, we get

$$\begin{aligned} & \sum_{i=1}^n \langle \nabla_1 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_j}(x) \rangle \nabla_2^2 a_\zeta(x, y) \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_m} \right) + \sum_{i, \eta=1}^n \langle \nabla_1 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_j}(x) \rangle a_\eta(x, y) \frac{\partial}{\partial x_m}(y) \Gamma_{i\eta}^\zeta \\ & + \sum_{i=1}^n \left(\langle \nabla_2 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_m}(y) \rangle \nabla_1 \nabla_2 a_\zeta(x, y) \left(\frac{\partial}{\partial x_i}(y), \frac{\partial}{\partial x_j}(x) \right) + o(1) \right) \\ & = 0, \end{aligned}$$

where we have used (2.11), (2.14), (2.26), (2.27), (2.36) and (2.39). Letting y approach to x , we obtain

$$\nabla_1 \nabla_2 a_\zeta(x, x) \left(\frac{\partial}{\partial x_m}(x), \frac{\partial}{\partial x_j}(x) \right) = \nabla_2^2 a_\zeta(x, x) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_m} \right) + \sum_{\eta=1}^n a_\eta(x, x) \frac{\partial}{\partial x_m}(x) \Gamma_{j\eta}^\zeta, \quad (2.41)$$

where we have used (2.35). For $j = 1, \dots, n$, letting $\frac{\partial}{\partial x_j}(y)$ act on (2.40) with $z = x$, we get

$$\begin{aligned} & \sum_{i=1}^n \left(\langle \nabla_2 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_m}(y) \rangle \nabla_2^2 a_\zeta(x, y) \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right) + \langle \nabla_2 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_j}(y) \rangle \nabla_2^2 a_\zeta(x, y) \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_m} \right) \right) \\ & + \sum_{i, \eta=1}^n \left(\langle \nabla_2 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_m}(y) \rangle a_\eta(x, y) \frac{\partial}{\partial x_j}(y) \Gamma_{i\eta}^\zeta + \langle \nabla_2 \dot{\gamma}_{xy}^i(1), \frac{\partial}{\partial x_j}(y) \rangle a_\eta(x, y) \frac{\partial}{\partial x_m}(y) \Gamma_{i\eta}^\zeta \right) \\ & + o(1) = 0, \end{aligned}$$

where we have used (2.11), (2.14), (2.26), (2.27), (2.36) and (2.39). Letting y approach to x , we obtain

$$\begin{aligned} & \nabla_2^2 a_\zeta(x, x) \left(\frac{\partial}{\partial x_m}, \frac{\partial}{\partial x_j} \right) + \nabla_2^2 a_\zeta(x, x) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_m} \right) \\ & + \sum_{\eta=1}^n a_\eta(x, x) \left(\frac{\partial}{\partial x_j}(x) \Gamma_{m\eta}^\zeta + \frac{\partial}{\partial x_m}(x) \Gamma_{j\eta}^\zeta \right) = 0, \end{aligned}$$

where we have used (2.35). Since $\nabla_2^2 a_\eta(x, x)$ is the Hessian of the function $a_\eta(x, \cdot)$ with respect to the second argument at x , it is a symmetric tensor of order $(0, 2)$. Therefore, we have

$$\nabla_2^2 a_\zeta(x, x) \left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_m} \right) = -\frac{1}{2} \sum_{\eta=1}^n a_\eta(x, x) \left(\frac{\partial}{\partial x_j}(x) \Gamma_{m\eta}^\zeta + \frac{\partial}{\partial x_m}(x) \Gamma_{j\eta}^\zeta \right). \quad (2.42)$$

Recalling (2.33) and inserting the above identity into (2.41), we can get (2.31) with $X = \frac{\partial}{\partial x_m}$ and $V = \frac{\partial}{\partial x_j}$, where we have used the expression of the curvature tensor in local coordinates (See [31, p. 41]), the property of the curvature tensor (See [31, p. 33]), (2.26), and the definition of $a_j(x, x)$ with $j = 1, \dots, n$. It is easy to check that $\nabla_{X(x)}(d_x(L_x.F(x))V(x))$ with $X, V \in TM$ is multi-linear with respect to X and V . Thus, the proof of (2.31) is completed.

To obtain (2.32), for any $y \in O$, we apply (2.22) and get

$$\begin{aligned} & \nabla_1 \nabla_2 \rho^2(x, y) (F(y), \frac{\partial}{\partial x_j}(x)) \\ &= -\nabla_1^2 \rho^2(x, y) (L_{yx} F(y), \frac{\partial}{\partial x_j}(x)) - \left\langle \nabla_1 \rho^2(x, y), \nabla_{\frac{\partial}{\partial x_j}(x)} (L_y \cdot F(y)) \right\rangle. \end{aligned}$$

Letting $\frac{\partial}{\partial x_k}(y)$ ($1 \leq k \leq n$) act on the above identity, via (2.11), we can get

$$\begin{aligned} & \nabla_1 \nabla_2^2 \rho^2(x, y) (F(y), \frac{\partial}{\partial x_k}(y), \frac{\partial}{\partial x_j}(x)) + \nabla_1 \nabla_2 \rho^2(x, y) (\nabla_{\frac{\partial}{\partial x_k}(y)} F, \frac{\partial}{\partial x_j}(x)) \\ &+ \nabla_2 \nabla_1^2 \rho^2(x, y) (L_{yx} F(y), \frac{\partial}{\partial x_j}(x), \frac{\partial}{\partial x_k}(y)) + \nabla_1^2 \rho^2(x, y) \left(d_y (L_{yx} F(y)) \frac{\partial}{\partial x_k}(y), \frac{\partial}{\partial x_j}(x) \right) \\ &= -\nabla_2 \nabla_1 \rho^2(x, y) \left(\nabla_{\frac{\partial}{\partial x_j}(x)} (L_y \cdot F(y)), \frac{\partial}{\partial x_k}(y) \right) \\ &- \left\langle \nabla_1 \rho^2(x, y), \nabla_{\frac{\partial}{\partial x_j}(x)} \left(d_y (L_y \cdot F(y)) \frac{\partial}{\partial x_k}(y) \right) \right\rangle. \end{aligned}$$

Moreover, letting $\frac{\partial}{\partial x_k}(y)$ act on the above identity, as y approaches to x , we can get

$$\begin{aligned} & \nabla_1 \nabla_2^3 \rho^2(x, x) (F, \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_j}) + \nabla_1 \nabla_2 \rho^2(x, x) \left(\nabla_{\frac{\partial}{\partial x_k}(x)} \nabla_{\frac{\partial}{\partial x_k}} F, \frac{\partial}{\partial x_j}(x) \right) \\ &+ \nabla_2^2 \nabla_1^2 \rho^2(x, x) (F, \frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_k}) \\ &+ \nabla_1^2 \rho^2(x, x) \left(d_y \left(d_y (L_{yx} F(y)) \frac{\partial}{\partial x_k}(y) \right) \frac{\partial}{\partial x_k}(y) \Big|_{y=x}, \frac{\partial}{\partial x_j}(x) \right) \\ &= -2 \nabla_2 \nabla_1 \rho^2(x, x) \left(\nabla_{\frac{\partial}{\partial x_j}(x)} \left(d_y (L_y \cdot F(y)) \Big|_{y=x} \frac{\partial}{\partial x_k}(x) \right), \frac{\partial}{\partial x_k}(x) \right), \end{aligned} \tag{2.43}$$

where (2.17), (2.18) and (2.24) are used.

We claim that

$$\begin{aligned} & \nabla_1 \nabla_2 \rho^2(x, x) \left(\nabla_{\frac{\partial}{\partial x_k}(x)} \nabla_{\frac{\partial}{\partial x_k}} F, \frac{\partial}{\partial x_j}(x) \right) \\ &+ \nabla_1^2 \rho^2(x, x) \left(d_y \left(d_y (L_{yx} F(y)) \frac{\partial}{\partial x_k}(y) \right) \frac{\partial}{\partial x_k}(y) \Big|_{y=x}, \frac{\partial}{\partial x_j}(x) \right) = 0. \end{aligned} \tag{2.44}$$

In fact, by (2.28), we may rewrite $F(y)$ as $F(y) = L_{yy} F(y) = \sum_{\eta=1}^n a_\eta(y, y) \frac{\partial}{\partial x_\eta}(y)$, and compute $\nabla_{\frac{\partial}{\partial x_k}(y)} F$ as follows:

$$\begin{aligned} \nabla_{\frac{\partial}{\partial x_k}(y)} F &= \sum_{\eta=1}^n \left(\langle \nabla_1 a_\eta(y, y), \frac{\partial}{\partial x_k}(y) \rangle + \langle \nabla_2 a_\eta(y, y), \frac{\partial}{\partial x_k}(y) \rangle \right) \frac{\partial}{\partial x_\eta}(y) \\ &+ \sum_{\eta, \zeta=1}^n a_\eta(y, y) \Gamma_{k\eta}^\zeta \frac{\partial}{\partial x_\zeta}(y). \end{aligned}$$

Applying Lemma 2.2, and by (2.26) and (2.27), we have

$$\begin{aligned} & \nabla_{\frac{\partial}{\partial x_k}(y)} \nabla_{\frac{\partial}{\partial x_k}} F \\ &= \sum_{\eta=1}^n \left(\nabla_1^2 a_\eta(y, y) \left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_k} \right) + 2 \nabla_1 \nabla_2 a_\eta(y, y) \left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_k} \right) \right. \\ &\quad \left. + \nabla_2^2 a_\eta(y, y) \left(\frac{\partial}{\partial x_k}, \frac{\partial}{\partial x_k} \right) \right) \frac{\partial}{\partial x_\eta}(y) + \sum_{\eta, \zeta=1}^n a_\eta(y, y) \frac{\partial}{\partial x_k}(y) (\Gamma_{k\eta}^\zeta) \frac{\partial}{\partial x_\zeta}(y) + o(1), \end{aligned} \tag{2.45}$$

where $\lim_{y \rightarrow x} o(1) = 0$. On the other hand, to figure out $d_y \left(d_y (L_{yx} F(y)) \frac{\partial}{\partial x_k}(y) \right) \frac{\partial}{\partial x_k}(y)$ with y closing to x , we choose a geodesic $\gamma(\cdot)$ on M satisfying $\gamma(0) = y$ and $\dot{\gamma}(0) = \frac{\partial}{\partial x_k}(y)$. Then,

$$d_y (L_{yx} F(y)) \frac{\partial}{\partial x_k}(y) = \frac{\partial}{\partial s} \Big|_0 \sum_{i=1}^n a_i(\gamma(s), x) \frac{\partial}{\partial x_i}(x) = \sum_{i=1}^n \langle \nabla_1 a_i(y, x), \frac{\partial}{\partial x_k}(y) \rangle \frac{\partial}{\partial x_i}(x),$$

and

$$\begin{aligned} d_y \left(d_y (L_{yx} F(y)) \frac{\partial}{\partial x_k} (y) \right) \frac{\partial}{\partial x_k} (y) &= \frac{\partial}{\partial s} \Big|_0 \sum_{i=1}^n \langle \nabla_1 a_i(\gamma(s), x), \frac{\partial}{\partial x_k} (\gamma(s)) \rangle \frac{\partial}{\partial x_i} (x) \\ &= \sum_{i=1}^n \nabla_1^2 a_i(y, x) \left(\frac{\partial}{\partial x_k} (y), \frac{\partial}{\partial x_k} (y) \right) \frac{\partial}{\partial x_i} (x) + o(1). \end{aligned} \quad (2.46)$$

Inserting (2.46) and (2.45) (with y approaching to x) into the left hand side of (2.44), via Lemma 2.2, (2.41) and (2.42), we get (2.44).

Inserting (2.44) into (2.43), via Lemma 2.2 and (2.31), we can get (2.32) with $X = \frac{\partial}{\partial x_j}$ and $V = \frac{\partial}{\partial x_k}$. It is easy to check that $\nabla_2^3 \nabla_1 \rho^2(x, x)(X, F, V, W)$, $\nabla_2^2 \nabla_1^2 \rho^2(x, x)(X, F, V, W)$ and $R(X, V, F, W)(x)$ with $X, F, V, W \in TM$ are multi-linear with respect to X, F, V and W . Hence, the desired identity (2.32) follows. \square

In particular, when M is a two dimensional Riemannian manifold, we can employ Gauss-Bonnet Theorem to prove (2.31).

Lemma 2.4 *Assume that M is a two dimensional Riemannian manifold. Let $X, V, F \in TM$. Then, for any $x \in M$, it holds that*

$$\nabla_{X(x)} \left(d_x (L_x F(x)) V(x) \right) = \frac{1}{2} k(x) \left(\langle F(x), V_X^\perp(x) \rangle X(x) - \langle F(x), X(x) \rangle V_X^\perp(x) \right), \quad (2.47)$$

where $k(x)$ is the Gaussian curvature of M at the point x , and $V_X^\perp(x)$ is given by

$$V_X^\perp(x) \equiv \begin{cases} V(x) - \left\langle V(x), \frac{X(x)}{|X(x)|} \right\rangle \frac{X(x)}{|X(x)|}, & \text{if } X(x) \neq 0, \\ V(x), & \text{if } X(x) = 0. \end{cases}$$

Remark 2.1 *Assume that M is a two dimensional Riemannian manifold. If $X, Y \in TM$ are orthonormal at some point $x \in M$, then one has $k(x) = \sec(X(x), Y(x))$. Applying this relation and formula (2.30), one can easily check that (2.47) is consistent with (2.31).*

Proof of Lemma 2.4 If $X(x) = 0$ or $V(x) = 0$, (2.47) holds obviously.

If $X(x) \neq 0$, $V(x) \neq 0$ and $V_X^\perp(x) \neq 0$, set $e_1 \equiv \frac{X(x)}{|X(x)|}$ and $e_2 = \frac{V_X^\perp(x)}{|V_X^\perp(x)|}$. Let us compute $\nabla_{e_1}(d_x L_x F(x) e_2)$ firstly. Set $\gamma_i(s) \equiv \exp_x(s e_i)$ with $s \geq 0$ and $i = 1, 2$. For any $y \in M$ closing enough to x , by the definition of the differential of a map, we have

$$d_x(L_{xy} F(x)) e_2 = d_x(L_{xy} F(x)) \dot{\gamma}_2(0) = \frac{d}{ds} \Big|_0 L_{\gamma_2(s)y} F(\gamma_2(s)).$$

Using the definition of the covariant derivative in the sense of parallel translation (2.9) and the above identity, we see that

$$\begin{aligned} & \nabla_{e_1}(d_x(L_x F(x)) e_2) \\ &= \lim_{\tau \rightarrow 0^+} \frac{1}{\tau} \left(L_{\gamma_1(\tau)x} d_x(L_{x\gamma_1(\tau)} F(x)) e_2 - d_x(L_{xx} F(x)) e_2 \right) \\ &= \lim_{\tau, s \rightarrow 0^+} \frac{1}{\tau s} \left(L_{\gamma_1(\tau)x} \left(L_{\gamma_2(s)\gamma_1(\tau)} F(\gamma_2(s)) - L_{x\gamma_1(\tau)} F(x) \right) - \left(L_{\gamma_2(s)x} F(\gamma_2(s)) \right. \right. \\ & \quad \left. \left. - F(x) \right) \right) \\ &= \lim_{\tau, s \rightarrow 0^+} \frac{1}{\tau s} \left(L_{\gamma_1(\tau)x} L_{\gamma_2(s)\gamma_1(\tau)} F(\gamma_2(s)) - L_{\gamma_2(s)x} F(\gamma_2(s)) \right) \end{aligned} \quad (2.48)$$

It is well known that the parallel translation of a vector along a curve conserves its norm and the angle between this vector and the curve. We will use this property to compute

$$L_{\gamma_1(\tau)x}L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s)) - L_{\gamma_2(s)x}F(\gamma_2(s)).$$

Let $\gamma(\cdot)$ be the shortest geodesic connecting $\gamma_1(\tau)$ and $\gamma_2(s)$ with $\gamma(0) = \gamma_2(s)$ and $\gamma(1) = \gamma_1(\tau)$. Denote by $\Delta_{s\tau}$ the domain inside the geodesic triangle with vertexes $\gamma_1(\tau)$, $\gamma_2(s)$ and x . The above formula represents the difference of parallel translations of $F(\gamma_2(s))$ along different curves.

For $s \geq 0$, denote by $\beta(s)$ the angle between $F(\gamma_2(s))$ and $\dot{\gamma}_2(s)$. Then, the angle between $L_{\gamma_2(s)x}F(\gamma_2(s))$ and $\dot{\gamma}_2(0) = e_2$ is $\beta(s)$, while the angle between $L_{\gamma_2(s)x}F(\gamma_2(s))$ and e_1 is $\frac{\pi}{2} - \beta(s)$. Therefore,

$$L_{\gamma_2(s)x}F(\gamma_2(s)) = |F(\gamma_2(s))| \left(\cos\left(\frac{\pi}{2} - \beta(s)\right)e_1 + \cos\beta(s)e_2 \right). \quad (2.49)$$

Denote the internal angles of the geodesic triangle $\Delta_{s\tau}$ at the points $\gamma_1(\tau)$ and $\gamma_2(s)$ by α_1 and α_2 , respectively. The angle between $F(\gamma_2(s))$ and $\dot{\gamma}(0)$ is $\pi - \alpha_2 - \beta(s)$, so is the angle between $L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s))$ and $\dot{\gamma}(1)$. Then, the angle between $L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s))$ and $\dot{\gamma}_1(\tau)$ is $\pi - \alpha_2 - \alpha_1 - \beta(s)$. Thus, the angle between $L_{\gamma_1(\tau)x}L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s))$ and $\dot{\gamma}_1(0) = e_1$ is $\pi - \alpha_2 - \alpha_1 - \beta(s)$. Consequently, the angle between $L_{\gamma_1(\tau)x}L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s))$ and e_2 is $\alpha_1 + \alpha_2 + \beta(s) - \frac{\pi}{2}$, due to the fact that e_1 is perpendicular to e_2 . Therefore,

$$\begin{aligned} & L_{\gamma_1(\tau)x}L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s)) \\ &= |F(\gamma_2(s))| \left(\cos(\pi - \alpha_1 - \alpha_2 - \beta(s))e_1 + \cos(\alpha_1 + \alpha_2 + \beta(s) - \frac{\pi}{2})e_2 \right). \end{aligned} \quad (2.50)$$

The Gauss-Bonnet formula (ref. [11, p. 274]) gives

$$\alpha_1 + \alpha_2 = \iint_{\Delta_{s\tau}} k dA + \frac{\pi}{2},$$

where dA is the Riemannian volume element. By using (2.49), (2.50), the above identity and Taylor's theorem, we have

$$\begin{aligned} & L_{\gamma_1(\tau)x}L_{\gamma_2(s)\gamma_1(\tau)}F(\gamma_2(s)) - L_{\gamma_2(s)x}F(\gamma_2(s)) \\ &= |F(\gamma_2(s))| \iint_{\Delta_{s\tau}} k dA \left(\cos\beta(s)e_1 - \sin\beta(s)e_2 \right) + o\left(\iint_{\Delta_{s\tau}} k dA\right). \end{aligned} \quad (2.51)$$

We choose $\{s, \tau\}$ as the normal coordinates at the point x . When s and τ are small enough, the area of $\Delta_{s\tau}$, denoted by $A(\Delta_{s\tau})$, is given by

$$A(\Delta_{s\tau}) = \int_0^s \langle \exp_x^{-1} \gamma(\frac{\xi}{s}), e_1 \rangle d\xi.$$

Set $v(s, \tau) \equiv \exp_{\gamma_2(s)}^{-1} \gamma_1(\tau)$. Then, $\gamma(\zeta) = \exp_{\gamma_2(s)}(\zeta v(s, \tau))$ with $\zeta \in [0, 1]$ and

$$\begin{aligned}
& \lim_{s, \tau \rightarrow 0^+} \frac{1}{s\tau} A(\Delta_{s\tau}) \\
&= \lim_{s \rightarrow 0^+} \frac{1}{s} \int_0^s \left\langle d \exp_x^{-1} \Big|_{\exp_{\gamma_2(s)} \frac{\xi}{s} v(s, 0)} \circ d \exp_{\gamma_2(s)} \Big|_{\frac{\xi}{s} v(s, 0)} \frac{\xi}{s} (d \exp_{\gamma_2(s)}^{-1} \Big|_x e_1), e_1 \right\rangle d\xi \\
&\stackrel{\zeta=\xi/s}{=} \lim_{s \rightarrow 0^+} \int_0^1 \left\langle d \exp_x^{-1} \Big|_{\exp_{\gamma_2(s)} \zeta v(s, 0)} \circ d \exp_{\gamma_2(s)} \Big|_{\zeta v(s, 0)} \zeta (d \exp_{\gamma_2(s)}^{-1} \Big|_x e_1), e_1 \right\rangle d\zeta \\
&= \frac{1}{2},
\end{aligned}$$

where we have used Lemma 2.1 and the L'Hôpital's rule. Recalling (2.48) and (2.51), we have

$$\begin{aligned}
& \nabla_{e_1}(d_x L_x \cdot F(x) e_2) \\
&= \lim_{s, \tau \rightarrow 0^+} \frac{1}{2} \frac{1}{A(\Delta_{s\tau})} \left(|F(\gamma_2(s))| \int_{\Delta_{s\tau}} k dA \left(\cos \beta(s) e_1 - \cos \left(\frac{\pi}{2} - \beta(s) \right) e_2 \right) + o \left(\int_{\Delta_{s\tau}} k dA \right) \right) \\
&= \frac{1}{2} k(x) |F(x)| \left(\cos \beta(0) e_1 - \cos \left(\frac{\pi}{2} - \beta(0) \right) e_2 \right) \\
&= \frac{1}{2} k(x) \left(\left\langle F(x), \frac{V_X^\perp}{|V_X^\perp|} \right\rangle \frac{X}{|X|} - \left\langle F(x), \frac{X}{|X|} \right\rangle \frac{V_X^\perp}{|V_X^\perp|} \right),
\end{aligned}$$

which implies (2.47).

If $X \neq 0$, $V \neq 0$ and $V_X^\perp = 0$, V is parallel to X . Similar to (2.48), we can get $\nabla_{e_1}(d_x L_x \cdot F(x) e_1) = 0$, which implies $\nabla_X(d_x L_x \cdot F(x) V) = 0$. The proof of Lemma 2.4 is complete. \square

Remark 2.2 If $X(x) = e_1$ and $V(x) = e_2$ are orthonormal, the left hand side of (2.31) characterizes the limiting case of parallel translations of vector field F along different curves. More precisely, by setting $\gamma_i(s) = \exp_x(s e_i)$ with $s \geq 0$ and $i = 1, 2$, we can get (2.48).

3 Statement of the main results

We begin with the following assumptions (Actually, as we shall see later, for our **Problem I**, we only need the assumptions (C1) and (C2) below):

(C1) The maps $f(= f(t, x, u)) : [0, T] \times M \times U \rightarrow TM$ and $f^0(= f^0(t, x, u)) : [0, T] \times M \times U \rightarrow \mathbb{R}$ are measurable in t , continuous in u , and C^1 in x . Moreover, there exists a constant $L > 1$ such that,

$$\begin{aligned}
|f^0(s, x_1, u) - f^0(s, x_2, u)| &\leq L\rho(x_1, x_2), \\
|L_{x_1 x_2} f(s, x_1, u) - f(s, x_2, u)| &\leq L\rho(x_1, x_2), \\
|f^0(s, x_0, u)| &\leq L, \quad |f(s, x_0, u)| \leq L,
\end{aligned} \tag{3.1}$$

for all $s \in [0, T]$, $u \in U$, and $x_1, x_2 \in M$ with $\rho(x_1, x_2) \leq \min\{i(x_1), i(x_2)\}$, where $x_0 \in M$ is arbitrarily fixed.

(C2) The maps $f(t, x, u)$ and $f^0(t, x, u)$ are C^2 in x . Furthermore,

$$\begin{aligned} |\nabla_x f^0(t, x_1, u) - L_{x_2 x_1} \nabla_x f^0(t, x_2, u)| &\leq L\rho(x_1, x_2), \\ |\nabla_x f(t, x_1, u) - L_{x_2 x_1} \nabla_x f(t, x_2, u)| &\leq L\rho(x_1, x_2), \end{aligned} \quad (3.2)$$

for all $x_1, x_2 \in M$ with $\rho(x_1, x_2) < \min\{i(x_1), i(x_2)\}$ and $(t, u) \in [0, T] \times U$, where $\nabla_x f(s, \cdot, u)$ and $\nabla_x f^0(s, \cdot, u)$ are the covariant derivatives of $f(s, \cdot, u)$ and $f^0(s, \cdot, u)$ with respect to the state variable, and their norms are given by (2.5).

(C3) $U \subset \mathbb{R}^m$ is open. The maps $f(t, x, u)$ and $f^0(t, x, u)$ are C^2 in u . Furthermore,

$$\begin{aligned} |\nabla_u f^0(t, x, u_1) - \nabla_u f^0(t, x, u_2)| &\leq L|u_1 - u_2|, \\ |\nabla_u f(t, x, u_1) - \nabla_u f(t, x, u_2)| &\leq L|u_1 - u_2|, \end{aligned} \quad (3.3)$$

for all $u_1, u_2 \in U$ and $(t, x) \in [0, T] \times M$, where $\nabla_u f(s, x, \cdot)$ and $\nabla_u f^0(s, x, \cdot)$ are the derivatives of $f(s, x, \cdot)$ and $f^0(s, x, \cdot)$ with respect to the control variable.

In this paper, for $X \in TM$ (or $X \in T^*M$), we denote by \tilde{X} the dual covector (vector) of X . Denote by $H^\nu : [0, T] \times T^*M \times U \rightarrow \mathbb{R}$ the Hamiltonian function, defined by

$$H^\nu(t, x, p, u) \equiv p(f(t, x, u)) + \nu f^0(t, x, u), \quad \forall (t, x, p, u, \nu) \in [0, T] \times T^*M \times U \times \mathbb{R}^-, \quad (3.4)$$

with $\mathbb{R}^- \equiv (-\infty, 0]$. In particular, when $\nu = -1$, we denote by

$$H \equiv H^{-1} \quad (3.5)$$

for abbreviation.

In this section, we fix a control $\bar{u}(\cdot) \in \mathcal{U}_{ad}$, where \mathcal{U}_{ad} is given in (1.3). Let $\bar{y}(\cdot)$ be the solution to (1.1) associated to $\bar{u}(\cdot)$. For abbreviation, we denote

$$[t] \equiv (t, \bar{y}(t), \bar{u}(t)), \quad \forall t \in [0, T]. \quad (3.6)$$

Suppose that $\psi(t) \in T_{\bar{y}(t)}^*M$ is the solution to the following first order dual equation¹:

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} \psi = -\nabla_x f[t](\psi(t), \cdot) - \nu d_x f^0[t], & a.e. t \in [0, T], \\ \psi(T) = \psi_1 \end{cases} \quad (3.7)$$

with $d_x f^0$ denoting the exterior derivative of f^0 with respect to the state variable x , and $\nabla_x f[t](\psi(t), \cdot)$ ($t \in [0, T]$) being a tensor given by

$$\nabla_x f[t](\psi(t), X(\bar{y}(t))) \equiv \nabla_{X(\bar{y}(t))} f(t, \cdot, \bar{u}(t))(\psi(t)), \quad \forall X \in TM.$$

¹Our first order dual equation (3.7) which contains the Levi-Civita connection ∇ , is the same as the dual equation (12.30) in [1, p. 168], if they are rewritten in the local coordinates. Nevertheless, the explicit form (3.7) is convenient for us to derive further the second order necessary condition for optimal controls by introducing a suitable second order dual equation (i.e., (3.8) below).

3.1 Optimality conditions for systems without endpoint constraints

Let $\psi(\cdot)$ be the solution to (3.7) with $\nu = -1$ and $\psi_1 = 0$. Recalling (3.5), we put

$$\tilde{U}(t) \equiv \{u \in U; \quad H(t, \bar{y}(t), \psi(t), \bar{u}(t)) = H(t, \bar{y}(t), \psi(t), u)\}, \quad t \in [0, T].$$

Let $w(t) \in \mathcal{T}_2^0(\bar{y}(t))$ with $t \in [0, T]$, and satisfy the following equation

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} w + \mathcal{E}^{12}(\nabla_x f[t] \otimes w(t)) + \mathcal{E}^{12}(w(t) \otimes \nabla_x f[t]) \\ \quad + \nabla_x^2 H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - R(\tilde{\psi}(t), \cdot, f[t], \cdot) = 0, \quad t \in [0, T], \\ w(T) = 0, \end{cases} \quad (3.8)$$

where \mathcal{E}^{12} is defined in (2.16), the tensor $R(\tilde{\psi}(t), \cdot, f[t], \cdot)$ is given by

$$R(\tilde{\psi}(t), \cdot, f[t], \cdot)(X, Y) \equiv R(\tilde{\psi}(t), X, f[t], Y), \quad \forall X, Y \in TM,$$

and the tensor $\nabla_x^2 H(t, \bar{y}(t), \psi(t), \bar{u}(t))$ is given by

$$\nabla_x^2 H(t, \bar{y}(t), \psi(t), \bar{u}(t))(X, Y) \equiv \nabla_x^2 f[t](\psi(t), X, Y) - \nabla_x^2 f^0[t](X, Y), \quad \forall X, Y \in TM.$$

Recalling the tensor contraction map in Section 2.4, we introduce two tensors $\Phi(t) \in \mathcal{T}_0^2(\bar{y}(t))$ and $\Phi_1(t) \in \mathcal{T}_2^0(\bar{y}(t))$ with $t \in [0, T]$, which solve the following equations respectively

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} \Phi = \mathcal{E}^{21}(\nabla_x f[t] \otimes \Phi(t)), \quad t \in (0, T], \\ \Phi(0)(X, Y) = \langle X, Y \rangle, \quad \forall X, Y \in T_{y_0}^* M, \end{cases} \quad (3.9)$$

and

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} \Phi_1 = -\mathcal{E}^{12}(\Phi_1(t) \otimes \nabla_x f[t]), \quad t \in (0, T], \\ \Phi_1(0)(X, Y) = \langle X, Y \rangle, \quad \forall X, Y \in T_{y_0} M. \end{cases} \quad (3.10)$$

Our second order necessary conditions for optimal controls without endpoint constraints are stated as follows.

Theorem 3.1 *Assume that the assumptions (C1) and (C2) hold, and (U, \tilde{d}) is a separable metric space. Let $(\bar{y}(\cdot), \bar{u}(\cdot))$ be an optimal pair for **Problem I**. Then,*

$$H(t, \bar{y}(t), \psi(t), \bar{u}(t)) = \max_{u \in U} H(t, \bar{y}(t), \psi(t), u), \quad \text{a.e. } t \in [0, T], \quad (3.11)$$

where $\psi(\cdot)$ is the solution to the first order dual equation (3.7) with $\nu = -1$ and $\psi_1 = 0$. Furthermore, for any $u(\cdot) \in \mathcal{U}_{ad}$ with $u(t) \in \tilde{U}(t)$, a.e. $t \in [0, T]$, it holds that

$$\begin{aligned} & \int_0^T \int_0^t \left\{ \nabla_x H(t, \bar{y}(t), \psi(t), u(t)) - \nabla_x H(t, \bar{y}(t), \psi(t), \bar{u}(t)), \right. \\ & \quad \mathcal{E}^{21}(\mathcal{E}^{21}(\Phi(t) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} \Phi_1(s)) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)}(f(s, \bar{y}(s), u(s)) - f[s])) \\ & \quad + \frac{1}{2}(w(t) + w(t)^\top)(f(t, \bar{y}(t), u(t)) - f[t], \mathcal{E}^{21}(\mathcal{E}^{21}(\Phi(t) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} \Phi_1(s)) \otimes \\ & \quad \left. L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)}(f(s, \bar{y}(s), u(s)) - f[s])) \right\} ds dt \\ & \leq 0, \end{aligned} \quad (3.12)$$

where the 2-form $w^\top(t)$ is the transpose of $w(t)$, defined by $w^\top(t)(X, V) \equiv w(t)(V, X)$ for any $X, V \in T_{\bar{y}(t)}M$, \mathcal{E}^{21} is defined in (2.16), and the tensor $\nabla_x H(t, \bar{y}(t), \psi(t), u)$ (with $u \in U$) is given by

$$\nabla_x H(t, \bar{y}(t), \psi(t), u)(X) \equiv \nabla_x f(t, \bar{y}(t), u)(\psi(t), X) - d_x f^0(t, \bar{y}(t), u)(X), \quad \forall X \in TM. \quad (3.13)$$

Moreover, if U is a Polish space, then the following pointwise condition holds:

$$\begin{aligned} & \frac{1}{2} \left(w(t) + w^\top(t) \right) \left(f[t] - f(t, \bar{y}(t), v), f[t] - f(t, \bar{y}(t), v) \right) \\ & + \left(\nabla_x H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - \nabla_x H(t, \bar{y}(t), \psi(t), v) \right) \left(f[t] - f(t, \bar{y}(t), v) \right) \\ & \leq 0, \quad \forall v \in \tilde{U}(t), \quad \text{a.e. } t \in [0, T]. \end{aligned} \quad (3.14)$$

Remark 3.1 The explicit dependence of the second order dual variable $w(\cdot)$ (involved in equation (3.8)) on the curvature tensor $R(\tilde{\psi}, \cdot, f, \cdot)$ reveals that the second order necessary condition (3.14) depends on the curvature. When M is \mathbb{R}^n , the curvature is zero everywhere, and for this special case, Theorem 3.1 coincides with [26, Theorem 4.3].

Denote by $d(\cdot, \cdot)$ the Ekeland metric over the admissible control space \mathcal{U}_{ad} , given by

$$d(u_1(\cdot), u_2(\cdot)) = |\{t \in [0, T]; u_1(t) \neq u_2(t)\}|, \quad \forall u_1(\cdot), u_2(\cdot) \in \mathcal{U}_{ad}.$$

It is well known that (\mathcal{U}_{ad}, d) is a complete metric space.

Our second order sufficient condition for locally optimal controls for **Problem I** is given as follows:

Theorem 3.2 Assume that (C1) and (C2) hold, and (U, \tilde{d}) is a separable metric space. If an admissible pair $(\bar{y}(\cdot), \bar{u}(\cdot))$ for **Problem I** satisfies (3.11), and there exist an $\epsilon_0 > 0$ and a $\beta > 0$ satisfying

$$\begin{aligned} & \int_0^T \int_0^t \left\{ \left\langle \nabla_x H(t, \bar{y}(t), \psi(t), u(t)) - \nabla_x H(t, \bar{y}(t), \psi(t), \bar{u}(t)), \right. \right. \\ & \quad \mathcal{E}^{21} \left(\mathcal{E}^{21} \left(\Phi(t) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} \Phi_1(s) \right) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} (f(s, \bar{y}(s), u(s)) - f[s]) \right) \Big\rangle \\ & \quad + \frac{1}{2} \left(w(t) + w(t)^\top \right) \left(f(t, \bar{y}(t), u(t)) - f[t], \mathcal{E}^{21} \left(\mathcal{E}^{21} \left(\Phi(t) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} \Phi_1(s) \right) \otimes \right. \right. \\ & \quad \left. \left. L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} (f(s, \bar{y}(s), u(s)) - f[s]) \right) \right) \Big\rangle ds dt \\ & \leq -\beta d(u(\cdot), \bar{u}(\cdot))^2, \quad \forall u(\cdot) \in \mathcal{U}_{ad} \text{ with } d(u(\cdot), \bar{u}(\cdot)) < \epsilon_0, \end{aligned} \quad (3.15)$$

then, one can find an $\epsilon_1 > 0$ such that

$$J(v(\cdot)) \geq J(\bar{u}(\cdot)), \quad \forall v(\cdot) \in \mathcal{V} \equiv \{u(\cdot) \in \mathcal{U}_{ad}; \quad d(u(\cdot), \bar{u}(\cdot)) \leq \epsilon_1\}.$$

3.2 Optimality conditions for systems with endpoint constraints

Before stating the main results of this subsection, we introduce some more notations. Recalling (3.4), under the assumptions (C2) and (C3), we introduce the following two tensors:

$$\nabla_x^2 H^\nu(t, x, p, u)(X, Y) = \nabla_x^2 f(t, x, u)(p, X, Y) + \nu \nabla_x^2 f^0(t, x, u)(X, Y), \quad (3.16)$$

$$\nabla_u \nabla_x H^\nu(t, x, p, u)(X, \eta) = \frac{\partial}{\partial u} \left(\nabla_x f(t, x, u)(p, X) + \nu d_x f^0(t, x, u)(X) \right) \cdot \eta, \quad (3.17)$$

for all $(t, x, p, u, \nu) \in [0, T] \times T^*M \times U \times \mathbb{R}^-$ and $X, Y \in TM$ and $\eta \in \mathbb{R}^m$.

Assume that the condition (C1) hold and (U, \tilde{d}) is a separable metric space. If $(\bar{y}(\cdot), \bar{u}(\cdot))$ is a optimal pair for **Problem II**, then there exist a $\nu \leq 0$ and a $\psi_1 \in T_{\bar{y}(T)}^*M$ with $|\nu| + |\psi_1| > 0$, such that the following Pontryagin's type maximum principle holds ([1, p. 181, Theorem 12.10]):

$$H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t)) = \max_{u \in U} H^\nu(t, \bar{y}(t), \bar{\psi}(t), u), \quad a.e. t \in [0, T], \quad (3.18)$$

where the covector field $\bar{\psi}(\cdot)$ along $\bar{y}(\cdot)$ solves (3.7). Moreover, when $U \subset \mathbb{R}^m$ is open, and f and f^0 are C^1 in u , it is easy to see that (3.18) implies that

$$\left. \frac{\partial}{\partial u} \right|_{\bar{u}(t)} H^\nu(t, \bar{y}(t), \bar{\psi}(t), u) = 0, \quad a.e. t \in [0, T]. \quad (3.19)$$

For any $\xi(\cdot) \in L^2(0, T; \mathbb{R}^m)$, we denote by $V(t) \in T_{\bar{y}(t)}M$ ($t \in [0, T]$) the solution to the following first order variational equation:

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} V = \nabla_{V(t)} f(t, \cdot, \bar{u}(t)) + \nabla_u f[t](\cdot, \xi(t)), & a.e. t \in (0, T], \\ V(0) = V(T) = 0. \end{cases} \quad (3.20)$$

Based on the above first order necessary condition, we give below a second order necessary condition for optimal pairs of **Problem II**.

Theorem 3.3 *Assume that the conditions (C1), (C2) and (C3) hold, and $(\bar{y}(\cdot), \bar{u}(\cdot))$ is an optimal pair for **Problem II**. Assume that $U \subset \mathbb{R}^m$ is open, $\bar{u}(\cdot) \in L^2(0, T; U)$, and for any neighborhood \mathcal{O} of $\bar{u}(\cdot)$ in $L^2(0, T; \mathbb{R}^m)$, there exists a $v(\cdot) \in \mathcal{V}_{ad} \cap \mathcal{O}$. If (ν, ψ_1) is a unique pair (up to a positive factor) satisfying (3.19), then the following second order necessary condition holds:*

(i) *If $\nu < 0$ (the normal case), then*

$$\begin{aligned} & \int_0^T \left\{ \nabla_u^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(\xi(t), \xi(t)) + \nabla_x^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), V(t)) \right. \\ & \quad \left. + 2\nabla_u \nabla_x H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), \xi(t)) - R(\bar{\psi}(t), V(t), \dot{\bar{y}}(t), V(t)) \right\} dt \\ & \leq 0, \end{aligned} \quad (3.21)$$

for all $\xi(\cdot) \in L^2(0, T; \mathbb{R}^m)$ and $V(\cdot) \in T_{\bar{y}(\cdot)}M$ satisfying (3.20), where the tensors $\nabla_x^2 H^\nu$ and $\nabla_u \nabla_x H^\nu$ are given respectively by (3.16) and (3.17).

(ii) If $\nu = 0$ (the abnormal case), the left hand side of (3.21) is sign-definite on the following set:

$$\left\{ \xi(\cdot) \in L^2(0, T; \mathbb{R}^m); \int_0^T \left\{ \langle \nabla_x f^0[t], V(t) \rangle + \nabla_u f^0[t](\xi(t)) \right\} dt = 0, \right. \\ \left. \text{where } V(\cdot) \in T_{\bar{y}(\cdot)} M \text{ solves (3.20)} \right\}. \quad (3.22)$$

Remark 3.2 In optimal control theory, usually one may use the needle variation technique to deduce a pointwise necessary condition from the integral one, provided that the later still holds when needle variations are used as control perturbations (e.g. the proof of [25, Theorem 1.6, pp. 150–158] for the first order case, and the proof of [26, Theorem 4.3, p. 1455] for the second order case). However, in our case, (3.21) does not enjoy such a property, due to the hard restriction (3.20). Hence, it seems quite difficult to derive a pointwise condition from (3.21).

Our second order sufficient conditions for locally optimal controls for **Problem II** is stated as follows:

Theorem 3.4 Assume that the conditions (C1), (C2) and (C3) hold, $U \subset \mathbb{R}^m$ is open, and $(\bar{y}(\cdot), \bar{u}(\cdot))$ is an admissible pair for **Problem II** satisfying (3.18) for some $\nu < 0$ and $\psi_1 \in T_{y_1}^* M$. If there exist a $\gamma_0 > 0$ such that

$$\frac{\partial^2}{\partial u^2} \Big|_{\bar{u}(t)} H^\nu(t, \bar{y}(t), \bar{\psi}(t), u)(v, v) \leq -\gamma_0 |v|^2, \quad \forall v \in \mathbb{R}^m, \text{ a.e. } t \in [0, T],$$

and a $\gamma_1 > 0$ so that

$$\int_0^T \left\{ \nabla_u^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(\xi(t), \xi(t)) + \nabla_x^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), V(t)) \right. \\ \left. + 2\nabla_u \nabla_x H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), \xi(t)) - R(\bar{\psi}(t), V(t), \dot{\bar{y}}(t), V(t)) \right\} dt \\ \leq -\gamma_1 \|\xi\|_{L^2(0, T; \mathbb{R}^m)}^2, \quad (3.23)$$

for all $\xi \in L^2(0, T; \mathbb{R}^m)$ and $V(\cdot) \in T_{\bar{y}(\cdot)} M$ satisfying (3.20), then $\bar{u}(\cdot)$ is locally optimal in the $L^2(0, T; U)$ topology.

4 Two illustrative examples

In this section, we shall give two simple but illustrative examples to show the usefulness of our results in the last section.

The first two lines of (3.1) and (3.2) are essentially Lipschitz conditions, and they are the key assumptions for our main results. The following result provides a method which can be employed to verify them.

Lemma 4.1 *Let \mathcal{T} be a tensor on M . Then, the following two conditions are equivalent:*

- (i) *There exists a positive constant L such that $|\nabla \mathcal{T}| \leq L$;*
- (ii) *There exists a positive constant L such that $|L_{x_1 x_2} \mathcal{T}(x_1) - \mathcal{T}(x_2)| \leq L\rho(x_1, x_2)$, for all $x_1, x_2 \in M$ with $\rho(x_1, x_2) < \min\{i(x_1), i(x_2)\}$.*

Proof: Suppose $\mathcal{T} \in TM$. Assume that (i) holds. Take any $x_1, x_2 \in M$ with $0 < \rho(x_1, x_2) < \min\{i(x_1), i(x_2)\}$. Set $V = \frac{1}{\rho(x_1, x_2)} \exp_{x_2}^{-1} x_1$. Then, by Lemma 2.2, we have $|V| = 1$. Assume that $\{e_1, e_2, \dots, e_n\}$ (with $e_1 \equiv V$) is an orthonormal basis at $T_{x_2} M$. Set $\gamma(s) \equiv \exp_{x_2} sV$ for $s \in [0, \tau]$ with $\tau \equiv \rho(x_1, x_2)$. Then, γ is the shortest geodesic connecting x_1 and x_2 , and parameterized by the arc length. Let $\{d_1, \dots, d_n\}$ be the dual basis of $\{e_1, \dots, e_n\}$ at x_2 . Set

$$e_i(s) = L_{x_2 \gamma(s)} e_i, \quad d_i(s) = L_{x_2 \gamma(s)} d_i, \quad s \in [0, \tau], \quad i = 1, \dots, n.$$

Then, $\{e_1(s), \dots, e_n(s)\}$ is the orthonormal basis at $T_{\gamma(s)} M$, $\{d_1(s), \dots, d_n(s)\}$ is the orthonormal basis at $T_{\gamma(s)}^* M$ for $s \in [0, \tau]$, and they are dual to each other. Thus, we can express \mathcal{T} along the curve γ as

$$\mathcal{T}(\gamma(s)) = \sum_{i=1}^n a_i(s) e_i(s), \quad \forall s \in [0, \tau].$$

Take any $v = \sum_{k=1}^n v_k d_k \in T_{x_2}^* M$. By (i), (2.10), (2.7), the parallel translation of $e_i(\cdot)$ and $d_i(\cdot)$ along γ for $i = 1, \dots, n$ and the Mean Value Theorem, one can get

$$\begin{aligned} & \left| \mathcal{T}(x_2)(v) - L_{x_1 x_2} \mathcal{T}(x_1)(v) \right| = \left| \sum_{i=1}^n a_i(0) e_i \left(\sum_{k=1}^n v_k d_k \right) - \mathcal{T}(x_1) \left(L_{x_2 x_1} v \right) \right| \\ &= \left| \sum_{i=1}^n a_i(0) v_i - \sum_{i=1}^n a_i(\tau) e_i(\tau) \left(\sum_{k=1}^n v_k d_k(\tau) \right) \right| = \left| \sum_{i=1}^n (a_i(0) - a_i(\tau)) v_i \right| \\ &= \left| \tau \sum_{i=1}^n \dot{a}_i(\zeta) v_i \right| = \left| \rho(x_1, x_2) \sum_{i=1}^n \dot{a}_i(\zeta) e_i(\zeta) \left(\sum_{k=1}^n v_k d_k(\zeta) \right) \right| \\ &= \left| \rho(x_1, x_2) \nabla_{\dot{\gamma}(\zeta)} \mathcal{T} \left(L_{x_2 \gamma(\zeta)} v \right) \right| = \left| \rho(x_1, x_2) \nabla \mathcal{T}(\gamma(\zeta)) \left(L_{x_2 \gamma(\zeta)} v, \dot{\gamma}(\zeta) \right) \right| \\ &\leq L |v| \rho(x_1, x_2), \end{aligned}$$

where $\zeta \in [0, 1]$, and we get (ii). Conversely, if (ii) holds, we can apply (2.9) to get (i).

When \mathcal{T} is a tensor of general order, one can apply (2.7) and the same method as above to get the equivalence of (i) and (ii). \square

By means of the above lemma, we give below an example which satisfies the assumptions (C1) and (C2).

Example 4.1 Consider the hyperbolic surface under the hyperboloid model

$$H_R^2 \equiv \{(x_1, x_2, x_3) \in \mathbb{R}^3; x_3^2 - (x_1^2 + x_2^2) = R^2, x_3 > 0\},$$

endowed with the metric $h = \iota^* m$, where $\iota : H_R^2 \rightarrow \mathbb{R}^3$ is the inclusion, m is the Minkowski metric $m = dx_1^2 + dx_2^2 - dx_3^2$, and h is the pullback metric. H_R^2 is non-compact, and its Gaussian curvature is -1 . For the details of this model, we refer to [24, Proposition 3.5, p. 38]. Let (H_R^2, Ψ) be the local coordinates with

$$\Psi(x_1, x_2) = \left(x_1, x_2, \sqrt{R^2 + x_1^2 + x_2^2}\right), \quad (x_1, x_2) \in \mathbb{R}^2.$$

Let $U = \{1, 2, 3, 4\}$ with the discrete metric. Denote

$$f(x, u) = u^3 e^{-(R^2 + x_1^2 + x_2^2)} \left(x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}\right), \quad f^0(x, u) = u^2 e^{-(R^2 + x_1^2 + x_2^2)},$$

for $x = (x_1, x_2) \in \mathbb{R}^2$ and $u \in U$. We consider the optimal control problem (1.1) and (1.2) on H_R^2 , with f and f^0 given above. Then, the conditions (C1) and (C2) are satisfied.

Proof: It is easy to check that the third line of (3.1) hold. We will apply Lemma 4.1 to check the first two inequalities of (3.1) and (3.2). By a direct computation, the metric h can be expressed in the local coordinates (x_1, x_2) by

$$h = \left(1 - \frac{x_1^2}{R^2 + x_1^2 + x_2^2}\right) dx_1 \otimes dx_1 - \frac{x_1 x_2}{R^2 + x_1^2 + x_2^2} (dx_1 \otimes dx_2 + dx_2 \otimes dx_1) + \left(1 - \frac{x_2^2}{R^2 + x_1^2 + x_2^2}\right) dx_2 \otimes dx_2.$$

Hence, ∇f and $\nabla^2 f$ in the coordinates $x = (x_1, x_2) \in \mathbb{R}^2$ are given respectively as follows:

$$\begin{aligned} \nabla_x f(x, u) = & u^3 e^{-(R^2 + x_1^2 + x_2^2)} \left((-2x_1 x_2 - \frac{x_1 x_2}{R^2}) \frac{\partial}{\partial x_1} \otimes dx_1 + (-2x_2^2 + 1 + \frac{x_1^2}{R^2}) \frac{\partial}{\partial x_1} \otimes dx_2 \right. \\ & \left. + (2x_1^2 - \frac{x_2^2}{R^2} - 1) \frac{\partial}{\partial x_2} \otimes dx_1 + (2x_1 x_2 + \frac{x_1 x_2}{R^2}) \frac{\partial}{\partial x_2} \otimes dx_2 \right), \end{aligned} \quad (4.1)$$

$$\nabla_x^2 f(x, u) = u^3 \sum_{i,j,l=1}^2 a_{ijl}(x_1, x_2) \frac{\partial}{\partial x_i} \otimes dx_j \otimes dx_l, \quad (4.2)$$

where $a_{ijl}(x_1, x_2) = e^{-(R^2 + x_1^2 + x_2^2)} (R^2 + x_1^2 + x_2^2)^{-1} p_{ijl}(x_1, x_2)$ with $p_{ijl}(x_1, x_2)$ being a polynomial of x_1 and x_2 for $i, j, l = 1, 2$. We also obtain that, for any tangent vector field $Y = a_1 \frac{\partial}{\partial x_1} + a_2 \frac{\partial}{\partial x_2}$ and any cotangent vector field $\eta = b_1 dx_1 + b_2 dx_2$,

$$\begin{aligned} |Y| \leq 1 \quad & \text{if and only if} \quad a_1^2 + a_2^2 \leq \frac{R^2 + x_1^2 + x_2^2}{R^2}, \\ |\eta| \leq 1 \quad & \text{if and only if} \quad b_1^2 + b_2^2 \leq \frac{R^2 + x_1^2 + x_2^2}{R^2}, \end{aligned} \quad (4.3)$$

where we have used the definition (2.5). Furthermore, for any $x = (x_1, x_2)$, by using (4.1), (4.2) and (4.3), we get

$$\begin{aligned}
|\nabla_x f(x, u)| &\equiv \sup\{\nabla_x f(x, u)(\eta, Y); \eta \in T_x^*M, Y \in T_xM, |Y| \leq 1, |\eta| \leq 1\} \\
&= u^3 \sup\left\{e^{-(R^2+x_1^2+x_2^2)}(a_1 \ a_2) \begin{pmatrix} -2x_1x_2 - \frac{x_1x_2}{R^2} & -2x_2^2 + 1 + \frac{x_1^2}{R^2} \\ 2x_1^2 - \frac{x_2^2}{R^2} - 1 & 2x_1x_2 + \frac{x_1x_2}{R^2} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}; \right. \\
&\quad \left. \max\{a_1^2 + a_2^2, b_1^2 + b_2^2\} \leq \frac{R^2 + x_1^2 + x_2^2}{R^2} \right\} \\
&\leq L
\end{aligned}$$

and

$$\begin{aligned}
|\nabla_x^2 f(x, u)| &\equiv \sup\left\{\nabla_x^2 f(x, u)(\eta, Y, Z); \eta \in T_x^*M, Y, Z \in T_xM, |\eta| \leq 1, |Y| \leq 1, |Z| \leq 1\right\} \\
&= u^3 \sup\left\{\sum_{i,j,l=1}^2 a_{ijl} \frac{\partial}{\partial x_i} \otimes dx_j \otimes dx_l \sum_{k,m,n=1}^2 (a_k dx_k, b_m \frac{\partial}{\partial x_m}, c_n \frac{\partial}{\partial x_n}); \right. \\
&\quad \left. \max\{a_1^2 + a_2^2, b_1^2 + b_2^2, c_1^2 + c_2^2\} \leq \frac{R^2 + x_1^2 + x_2^2}{R^2} \right\} \\
&= u^3 \sup\left\{\sum_{i,j,l=1}^2 a_{ijl} a_i b_j c_l; \max\{a_1^2 + a_2^2, b_1^2 + b_2^2, c_1^2 + c_2^2\} \leq \frac{R^2 + x_1^2 + x_2^2}{R^2} \right\} \\
&\leq L
\end{aligned}$$

for some positive constant L . By applying Lemma 4.1, we obtain the second inequalities of (3.1) and (3.2). Using the same way, one can get the first inequalities of (3.1) and (3.2). \square

By means of Pontryagin-type maximum principle, Sussmann ([34, Example 5.10.2]) gave a new proof of the classical result that the locally shortest curve connecting two fixed points on a Riemannian manifold must be a geodesic. In the following example, we shall apply our second order necessary condition for optimal controls with endpoint constraints (Theorem 3.3) to recover the well-known second variation of energy:

Example 4.2 *Given any two points $y_0, y_1 \in M$, assume that $\bar{y} : [0, T] \rightarrow M$ is a smooth curve with $\bar{y}(0) = y_0$, $\bar{y}(T) = y_1$ and constant speed $|\dot{\bar{y}}(t)| \equiv C > 0$ (for all $t \in [0, T]$). If $\bar{y}(\cdot)$ is a locally shortest curve connecting y_0 and y_1 , then it must be a geodesic. Moreover, if $\bar{y}(\cdot)$ is a geodesic, the following second variation of energy holds (see [31, p. 159]):*

$$\int_0^T \left\{ |\nabla_{\dot{\bar{y}}(t)} V|^2 + R(\dot{\bar{y}}(t), V(\bar{y}(t)), \dot{\bar{y}}(t), V(\bar{y}(t))) \right\} dt \geq 0 \quad (4.4)$$

for any $V \in TM$ with $V(y_0) = 0$ and $V(y_1) = 0$.

Proof According to the completeness of (M, g) , given any bounded domain $D \subset M$ such that $\bar{y}(t) \in D$ for all $t \in [0, T]$, one can find $f_1, \dots, f_m \in TM$ ($m \in \mathbb{N}$) with compact supports, such that

$$\text{span}\{f_1|_{\bar{D}}, \dots, f_m|_{\bar{D}}\} = \{X|_{\bar{D}}; X \in TM\}. \quad (4.5)$$

Indeed, by the completeness of (M, g) and the Hopf-Rinow Theorem (see [31, Theorem 16, p. 137]), we see that \bar{D} is compact. Hence, by [36, Lemma 1, p. 52] or [12, Theorem 3.7, p. 72], there exist a $\delta > 0$ and $x_1, \dots, x_l \in \bar{D}$ ($l \in \mathbb{N}$) such that, for all $x \in \bar{D}$, the map $\exp_x : B(O, \delta) \subset T_x M \rightarrow B_x(\delta)$ is diffeomorphic and $\cup_{i=1}^l B_{x_i}(\delta/2) \supset \bar{D}$. Therefore, for each $i = 1, \dots, l$, we can define vector fields $\{f_i^j(\cdot)\}_{j=1}^n$ on $B_{x_i}(\delta)$ as a basis for TM restricted to $B_{x_i}(\delta)$ as follows:

$$f_i^j(y) = d\exp_{x_i}|_{\exp_{x_i}^{-1}y} e_i^j, \quad \forall y \in B_{x_i}(\delta),$$

where $\{e_i^j\}_{j=1}^n$ is a basis at x_i . Then, we extend $\{f_i^j(\cdot)\}_{j=1}^n$ from $B_{x_i}(\delta)$ to M smoothly, which are denoted by $\{\tilde{f}_i^j\}_{j=1}^n$, such that $f_i^j|_{B_{x_i}(\delta/2)} = \tilde{f}_i^j|_{B_{x_i}(\delta/2)}$ and $\tilde{f}_i^j|_{M \setminus B_{x_i}(\delta)} = 0$, for $j = 1, \dots, n$. Thus, $\{\tilde{f}_i^j\}$ ($i = 1, \dots, l, j = 1, \dots, n$) can express linearly any vector field restricted to \bar{D} .

According to (4.5), we see that

$$\begin{cases} \dot{y}(t) = \sum_{i=1}^m u_i(t) f_i(y(t)), \text{ a.e. } t \in [0, T], \\ y(0) = y_0, \quad y(T) = y_1, \end{cases} \quad (4.6)$$

with $u(t) = (u_1(t), \dots, u_m(t)) \in \mathbb{R}^m$ a.e. $t \in [0, T]$, describes all absolutely continuous curves contained in D , with endpoints y_0 and y_1 . In order to apply our optimal control results to the above shortest curve problem, we need to seek an appropriate cost functional. To this end, set

$$\Omega = \{y : [0, T] \rightarrow M; y(0) = y_0, y(T) = y_1, \text{ and } y(\cdot) \text{ is absolutely continuous}\}.$$

Denote by $L(y)$ the length of $y \in \Omega$. [31, Proposition 17, p. 126] says that, if $\bar{y}(\cdot)$ is a minimizer of L in Ω , then it also minimizes $E(y) \equiv \frac{1}{2} \int_0^T |\dot{y}(t)|^2 dt$ within Ω . Therefore, we define the cost functional as follows:

$$J(u(\cdot)) = \frac{1}{2} \int_0^T |\dot{y}_u(t)|^2 dt = \frac{1}{2} \int_0^T \left| \sum_{i=1}^m u_i(t) f_i(y_u(t)) \right|^2 dt, \quad (4.7)$$

where $y_u(\cdot)$ is the solution to (4.6) corresponding to the control $u(\cdot) \in \mathcal{U} \equiv \{u : [0, T] \rightarrow \mathbb{R}^m; u(\cdot) \text{ is measurable, and } y_u(0) = y_0, y_u(T) = y_1\}$.

Hence, we choose (4.7) to be the desired cost functional. The corresponding Hamilton function is

$$H^\nu(t, y, p, u) \equiv p \left(\sum_{i=1}^m u_i f_i(y) \right) + \frac{1}{2} \nu \left| \sum_{i=1}^m u_i f_i(y) \right|^2, \quad (4.8)$$

for $(t, y, p, u, \nu) \in [0, T] \times T^*M \times \mathbb{R}^m \times \mathbb{R}^-$.

Suppose that $(\bar{y}(\cdot), \bar{u}(\cdot))$ is optimal for problem (4.6) and (1.6) with J given in (4.7). Then $\bar{y}(\cdot)$ is the locally shortest curve connecting y_0 and y_1 . Applying Theorem 3.3 to this specific problem, we obtain that

$$\begin{aligned} & \bar{\psi}(t) \left(\sum_{i=1}^m \bar{u}_i(t) f_i(\bar{y}(t)) \right) + \frac{\nu}{2} \left| \sum_{i=1}^m \bar{u}_i(t) f_i(\bar{y}(t)) \right|^2 \\ &= \max_{u=(u_1, \dots, u_m) \in \mathbb{R}^m} \left\{ \bar{\psi}(t) \left(\sum_{i=1}^m u_i f_i(\bar{y}(t)) \right) + \frac{\nu}{2} \left| \sum_{i=1}^m u_i f_i(\bar{y}(t)) \right|^2 \right\}, \quad a.e. t \in [0, T], \end{aligned} \quad (4.9)$$

where the first order dual variable $\bar{\psi}$ solves

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} \bar{\psi} = - \sum_{i=1}^m \bar{u}_i(t) \nabla f_i(\bar{y}(t))(\bar{\psi}(t), \cdot) - \nu \sum_{i,j=1}^m \bar{u}_i(t) \bar{u}_j(t) \nabla f_i(\bar{y}(t))(\tilde{f}_j(\bar{y}(t)), \cdot), \\ \bar{\psi}(T) = \psi_1, \end{cases} \quad a.e. t \in [0, T], \quad (4.10)$$

with

$$|\nu| + |\psi_1| > 0, \quad \nu \leq 0, \quad (4.11)$$

and \tilde{f}_j ($j = 1, \dots, m$) being the dual covector of f_j . Clearly, (4.9) is the first order necessary condition for $(\bar{y}(\cdot), \bar{u}(\cdot))$.

By differentiating (4.9) with respect to the control variable at $\bar{u}(t)$ for almost every $t \in [0, T]$, we have

$$\bar{\psi}(t)(f_i(\bar{y}(t))) + \nu \langle f_i(\bar{y}(t)), \dot{\bar{y}}(t) \rangle = 0, \quad a.e. t \in [0, T] \quad (4.12)$$

for $i = 1, \dots, m$. According to (4.5), for any $V \in TM$ such that $V(y_0) = 0$ and $V(y_1) = 0$, we can find suitable real valued functions on $[0, T]$: $\xi_1(\cdot), \dots, \xi_m(\cdot)$, such that

$$\nabla_{\dot{\bar{y}}(t)} V - \sum_{i=1}^m \bar{u}_i(t) \nabla_{V(\bar{y}(t))} f_i = \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)), \quad t \in [0, T]. \quad (4.13)$$

Combining (4.12) and the above identity, we deduce that

$$\int_0^T \left\{ \bar{\psi}(t) \left(\nabla_{\dot{\bar{y}}(t)} V - \sum_{i=1}^m \bar{u}_i(t) \nabla_{V(\bar{y}(t))} f_i \right) + \nu \langle \nabla_{\dot{\bar{y}}(t)} V - \sum_{i=1}^m \bar{u}_i(t) \nabla_{V(\bar{y}(t))} f_i, \dot{\bar{y}}(t) \rangle \right\} dt = 0.$$

Using integration by parts to the above identity, via the dual equation (4.10), we can get

$$-\nu \int_0^T \langle V(\bar{y}(t)), \nabla_{\dot{\bar{y}}(t)} \dot{\bar{y}} \rangle dt = 0. \quad (4.14)$$

If $\nu = 0$, (4.12) implies $\bar{\psi}(t) = 0$ for all $t \in [0, T]$, which contradicts (4.11). Hence, $\nu < 0$. By the choice of V , (4.14) implies $\nabla_{\dot{\bar{y}}(t)} \dot{\bar{y}} = 0$ for all $t \in [0, T]$, which means that $\bar{y}(\cdot)$ is a geodesic.

From (4.12) we can get

$$\dot{\bar{y}}(t) = -\frac{1}{\nu} \tilde{\bar{\psi}}(t), \quad \forall t \in [0, T], \quad (4.15)$$

which implies (ν, ψ_1) is unique up to a positive factor. In order to deduce the second order necessary condition, we need to compute $\nabla_u^2 H^\nu$, $\nabla_u \nabla_x H^\nu$ and $\nabla_x^2 H^\nu$ at $(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))$ respectively.

Recalling (4.8), for any $1 \leq i, j \leq n$, we have

$$\frac{\partial^2}{\partial u_i \partial u_j} \Big|_{\bar{u}(t)} H^\nu(t, \bar{y}(t), \bar{\psi}(t), u) = \nu \langle f_i(\bar{y}(t), f_j(\bar{y}(t))) \rangle, \quad \forall t \in [0, T]. \quad (4.16)$$

Hence, for any $\xi \in L^2(0, T; \mathbb{R}^m)$ satisfying (3.20) (or (4.13)), we get

$$\nabla_u^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(\xi(t), \xi(t)) = \nu \left| \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)) \right|^2. \quad (4.17)$$

Recalling (3.16), (3.17) and (4.15), we obtain that, via the first order variational equation (4.13),

$$\begin{aligned} & \nabla_x^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), V(t)) \\ &= \nu \left(|\nabla_{\dot{\bar{y}}(t)} V|^2 + \left| \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)) \right|^2 - 2 \langle \nabla_{\dot{\bar{y}}(t)} V, \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)) \rangle \right); \\ & \nabla_u \nabla_x H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), \xi(t)) = \nu \langle \nabla_{V(t)} \left(\sum_{j=1}^m \bar{u}_j(t) f_j \right), \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)) \rangle. \end{aligned}$$

Inserting (4.17) and the above two identities into (3.21), and via (4.13), (4.15) and (3.20), we end up with

$$\begin{aligned} 0 &\leq \int_0^T \left\{ |\nabla_{\dot{\bar{y}}(t)} V|^2 + R(\dot{\bar{y}}(t), V(t), \dot{\bar{y}}(t), V(t)) + 2 \langle \nabla_{V(t)} \left(\sum_{j=1}^m \bar{u}_j(t) f_j \right) \right. \\ &\quad \left. - \nabla_{\dot{\bar{y}}(t)} V, \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)) \rangle + 2 \left| \sum_{i=1}^m \xi_i(t) f_i(\bar{y}(t)) \right|^2 \right\} dt \\ &= \int_0^T \left\{ |\nabla_{\dot{\bar{y}}(t)} V|^2 + R(\dot{\bar{y}}(t), V(t), \dot{\bar{y}}(t), V(t)) \right\} dt, \end{aligned}$$

which implies (4.4). \square

Remark 4.1 *The Legendre condition ([1, Theorem 20.6, p. 300 and Proposition 20.11, p. 310]) says that the second order necessary condition of an optimal control $\bar{u}(\cdot)$ is :*

$$\frac{\partial^2}{\partial u^2} \Big|_{\bar{u}(t)} H^\nu(t, \bar{y}(t), \bar{\psi}(t), u)(v, v) \leq 0, \quad \forall v \in \mathbb{R}^m, \text{ a.e. } t \in [0, T].$$

Applying this condition to deal with the concrete problem in Example 4.2, instead of (4.4), one can only get

$$\nu \left| \sum_{i=1}^m v_i f_i(\bar{y}(t)) \right|^2 \leq 0, \quad \forall v = (v_1, \dots, v_m)^\top \in \mathbb{R}^m, \quad (4.18)$$

a.e. $t \in [0, T]$, where (4.16) is used. Clearly, the condition (4.18) is trivially correct. Hence, our second order necessary condition in Theorem 3.3 provides more information than that in [1]. In addition, from the inequality (4.18), it follows that the optimal control is not totally singular, i.e. $\bar{u}(\cdot)$ does not fulfills $\frac{\partial^2}{\partial u^2} \Big|_{\bar{u}(t)} H^\nu(t, \bar{y}(t), \bar{\psi}(t), u) = 0$ a.e. $t \in [0, T]$, and therefore, the Goh condition and the generalized Legendre condition do not work.

5 Variations of Trajectories

In this section, we shall give the first and second order variations of a trajectory of the control system (1.1), by employing respectively the needle variation in the case that the control set U is a metric space, and the classical variation in the case that U is an open subset in \mathbb{R}^m .

5.1 Needle Variation

In this subsection, we assume that U is a metric space.

Proposition 5.1 *Assume that the conditions (C1) and (C2) hold. Fix any $\bar{u}(\cdot) \in \mathcal{U}_{ad}$, and let $\bar{y}(\cdot)$ be the corresponding solution to (1.1). For any $u(\cdot) \in \mathcal{U}_{ad}$, denote by $X_u(t), Y_u(t) \in T_{\bar{y}(t)}M$ respectively solutions to the following first and second order variational equations:*

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} X_u = \nabla_{X_u(t)} f(t, \cdot, \bar{u}(t)) + f(t, \bar{y}(t), u(t)) - f[t], & a.e. t \in (0, T], \\ X_u(0) = 0, \end{cases} \quad (5.1)$$

and

$$\begin{cases} \langle \nabla_{\dot{\bar{y}}(t)} Y_u, Z \rangle = \langle \nabla_{Y_u(t)} f(t, \cdot, \bar{u}(t)), Z \rangle - \frac{1}{2} R(Z, X_u(t), f[t], X_u(t)) \\ \quad + \langle \nabla_{X_u(t)} f(t, \cdot, u(t)) - \nabla_{X_u(t)} f(t, \cdot, \bar{u}(t)), Z \rangle \\ \quad + \frac{1}{2} \nabla_x^2 f[t](\tilde{Z}, X_u(t), X_u(t)), & a.e. t \in (0, T], \\ Y_u(0) = 0, \end{cases} \quad (5.2)$$

for any $Z \in TM$, where $[t]$ is given in (3.6). Then, as $d(u(\cdot), \bar{u}(\cdot)) \equiv \epsilon \rightarrow 0$,

$$|X_u(t)| = O(\epsilon), \quad |Y_u(t)| = O(\epsilon^2), \quad (5.3)$$

$$|V_u(t) - X_u(t)| = o(\epsilon), \quad |V_u(t) - X_u(t) - Y_u(t)| = o(\epsilon^2), \quad (5.4)$$

where

$$V_u(t) = \exp_{\bar{y}(t)}^{-1} y_u(t), \quad t \in [0, T], \quad (5.5)$$

and $y_u(\cdot)$ is the corresponding solution to (1.1) with the control $u(\cdot)$.

To prove Proposition 5.1, we need the following result.

Lemma 5.1 *Under the assumptions in Proposition 5.1, for any $u(\cdot) \in \mathcal{U}_{ad}$, it holds that*

$$\rho(y_u(\hat{t}), y_u(t)) \leq (1 + \rho(x_0, y_0))(e^{L\hat{t}} - e^{Lt}), \quad \forall 0 \leq t \leq \hat{t} \leq T, \quad (5.6)$$

where $x_0 \in M$ and $L > 0$ are given in assumption (C1). Furthermore, there exists a constant $\epsilon_1 > 0$ such that

$$\rho(\bar{y}(t), y_u(t)) \leq 2L(1 + \rho(x_0, y_0))e^{Lt} \left| [0, t] \cap \{\tau \in [0, T]; u(\tau) \neq \bar{u}(\tau)\} \right|, \quad (5.7)$$

for any $t \in [0, T]$ and $u(\cdot) \in \mathcal{U}_{ad}$ with $d(\bar{u}(\cdot), u(\cdot)) < \epsilon_1$.

Proof: The proof is divided into two parts.

Part I. In this part, we prove (5.6). Firstly, we claim that

$$|f(t, y, v)| \leq L(\rho(x_0, y) + 1) \quad (5.8)$$

holds for any $t \in [0, T]$, $y \in M$ and $v \in U$.

In fact, by the completeness of the manifold M , we can find the shortest geodesic γ connecting x_0 and y with $\gamma(0) = x_0$ and $\gamma(1) = y$. Let $0 = s_0 < s_1 < s_2 < \dots < s_{N-1} < s_N = 1$ (for some $N \in \mathbb{N}$) be such that $\rho(\gamma(s_j), \gamma(s_{j-1})) < \min\{i(\gamma(s_j)), i(\gamma(s_{j-1}))\}$ for $j = 1, 2, \dots, N$. By means of (2.8), the condition (C1), and the triangle inequality of the norm $|\cdot|$ yields

$$\begin{aligned} & |f(t, y, v)| \\ & \leq |L_{\gamma(s_{N-1})\gamma(s_N)}f(t, \gamma(s_{N-1}), v) - f(t, y, v)| + |f(t, \gamma(s_{N-1}), v)| \\ & \leq |L_{\gamma(s_{N-1})\gamma(s_N)}f(t, \gamma(s_{N-1}), v) - f(t, y, v)| \\ & \quad + |f(t, \gamma(s_{N-1}), v) - L_{\gamma(s_{N-2})\gamma(s_{N-1})}f(t, \gamma(s_{N-2}), v)| + |f(t, \gamma(s_{N-2}), v)| \\ & \leq \dots \\ & \leq \sum_{j=1}^N |L_{\gamma(s_{j-1})\gamma(s_j)}f(t, \gamma(s_{j-1}), v) - f(t, \gamma(s_j), v)| + |f(t, x_0, v)| \\ & \leq L \sum_{j=1}^N \rho(\gamma(s_{j-1}), \gamma(s_j)) + L = L(\rho(x_0, y) + 1), \end{aligned}$$

which implies (5.8).

Secondly, we estimate $|f(t, y_u(t), v)|$ for any $v \in U$ and $t \in [0, T]$. Let $\beta : [0, 1] \rightarrow M$ be the shortest geodesic connecting $y_u(t)$ and y_0 and satisfy $\beta(0) = y_0$ and $\beta(1) = y_u(t)$. By the compactness, we can choose $\tau_0 = 0, \tau_1, \tau_2, \dots, \tau_l = 1$ for some $l \in \mathbb{N}$, such that $\tau_0 < \tau_1 < \dots < \tau_l$ and $\rho(\beta(\tau_i), \beta(\tau_{i+1})) < \min\{i(\beta(\tau_i)), i(\beta(\tau_{i+1}))\}$ for $i = 0, \dots, l-1$. Applying Lemma 2.2, the condition (C1), (5.8), and the triangle inequality of the norm $|\cdot|$, we obtain that

$$\begin{aligned} & |f(t, y_u(t), v)| \\ & \leq |f(t, y_u(t), v) - L_{\beta(\tau_{l-1})\beta(\tau_l)}f(t, \beta(\tau_{l-1}), v)| \\ & \quad + |f(t, \beta(\tau_{l-1}), v) - L_{\beta(\tau_{l-2})\beta(\tau_{l-1})}f(t, \beta(\tau_{l-2}), v)| + \dots + |f(t, y_0, v)| \quad (5.9) \\ & \leq L \sum_{i=1}^l \rho(\beta(\tau_i), \beta(\tau_{i-1})) + L(\rho(x_0, y_0) + 1) \\ & = L\rho(y_u(t), y_0) + L\rho(x_0, y_0) + L. \end{aligned}$$

Thirdly, we estimate $\rho(y_u(t), y_0)$ for any $t \in [0, T]$. Applying (5.9), we get

$$\rho(y_u(t), y_0) \leq \int_0^t |\dot{y}_u(s)| ds \leq L \int_0^t \rho(y_u(s), y_0) ds + Lt(\rho(y_0, x_0) + 1).$$

Applying the Gronwall inequality, we obtain

$$\rho(y_u(t), y_0) \leq (1 + \rho(x_0, y_0))(e^{Lt} - 1), \quad \forall t \in [0, T]. \quad (5.10)$$

Finally, for any $t, \hat{t} \in [0, T]$ with $t \leq \hat{t}$, applying (5.9) and (5.10), we have

$$\begin{aligned}\rho(y_u(t), y_u(\hat{t})) &\leq \int_t^{\hat{t}} |\dot{y}_u(s)| ds \\ &\leq L \int_t^{\hat{t}} (\rho(y_u(s), y_0) + \rho(x_0, y_0) + 1) ds \\ &\leq (1 + \rho(x_0, y_0))(e^{L\hat{t}} - e^{Lt}),\end{aligned}$$

which yields (5.6).

Part II. In this part, we prove (5.7). We will firstly show that (5.7) holds on a small interval, and then we extend it to the whole interval $[0, T]$. From (5.6), it follows that, for any $u(\cdot) \in \mathcal{U}_{ad}$, the corresponding trajectory $y_u(\cdot)$ is contained in the closed ball $\overline{B_{y_0}(R)}$ with $R \equiv (1 + \rho(x_0, y_0))(e^{LT} - 1)$, where $B_{y_0}(R)$ is defined in (2.3). Clearly $\overline{B_{y_0}(R)}$ is compact, because of the completeness of the manifold M and the Hopf-Rinow Theorem (see [31, Theorem 16, p. 137]). Therefore, there exists a $\delta > 0$ such that, for any $x \in \overline{B_{y_0}(R)}$, the map $\exp_x : B(O, \delta) \subset T_x M \rightarrow B_x(\delta)$ is diffeomorphic, due to [36, Lemma 1, p. 52] or a direct consequence of [12, Theorem 3.7, p. 72].

We claim that (5.7) holds on the interval $[0, t_1]$, where

$$t_1 = \frac{1}{L} \ln \frac{\delta + 2(1 + \rho(x_0, y_0))}{2(1 + \rho(x_0, y_0))}.$$

In fact, by the triangle inequality of the norm $|\cdot|$ and (5.6), for any $t \in [0, t_1]$, we have

$$\begin{aligned}\rho(y_u(t), \bar{y}(t)) &\leq \rho(y_u(t), y_0) + \rho(y_0, \bar{y}(t)) \\ &\leq 2(e^{Lt} - 1)(1 + \rho(x_0, y_0)) \\ &< \delta.\end{aligned}$$

Denote by $I_{u \neq \bar{u}}$ the indicator function of the set $\{\tau \in [0, T]; u(\tau) \neq \bar{u}(\tau)\}$. For $t \in [0, t_1]$, using the condition (C1), Lemma 2.2, (2.13), (2.6), (2.8), (5.6), (5.9), and (5.10), we get

$$\begin{aligned}&\frac{d}{dt} \rho^2(y_u(t), \bar{y}(t)) \\ &= \langle \nabla_1 \rho^2(y_u(t), \bar{y}(t)), \dot{y}_u(t) \rangle + \langle \nabla_2 \rho^2(y_u(t), \bar{y}(t)), \dot{\bar{y}}(t) \rangle \\ &= -2 \langle \exp_{y_u(t)}^{-1} \bar{y}(t), f(t, y_u(t), u(t)) \rangle - 2 \langle \exp_{\bar{y}(t)}^{-1} y_u(t), f[t] \rangle \\ &= -2 \langle \exp_{y_u(t)}^{-1} \bar{y}(t), f(t, y_u(t), u(t)) \rangle - 2 \langle \exp_{y_u(t)}^{-1} \bar{y}(t), f(t, y_u(t), \bar{u}(t)) - f(t, y_u(t), \bar{u}(t)) \rangle \\ &\quad - 2 \langle \exp_{\bar{y}(t)}^{-1} y_u(t), f[t] \rangle \\ &= 2 \langle \exp_{\bar{y}(t)}^{-1} y_u(t), L_{y_u(t)\bar{y}(t)} f(t, y_u(t), \bar{u}(t)) - f[t] \rangle \\ &\quad - 2 \langle \exp_{y_u(t)}^{-1} \bar{y}(t), f(t, y_u(t), u(t)) - f(t, y_u(t), \bar{u}(t)) \rangle I_{u \neq \bar{u}}(t) \\ &\leq 2L \rho(\bar{y}(t), y_u(t)) (\rho(\bar{y}(t), y_u(t)) + 2e^{Lt}(1 + \rho(x_0, y_0)) I_{u \neq \bar{u}}(t)),\end{aligned}$$

which leads to

$$\frac{d}{dt} \rho(y_u(t), y(t)) \leq L (\rho(\bar{y}(t), y_u(t)) + 2e^{Lt}(1 + \rho(x_0, y_0)) I_{u \neq \bar{u}}(t)). \quad (5.11)$$

Applying Gronwall's inequality to the above inequality, we get (5.7) on $[0, t_1]$. By the inequality (5.6) and the continuity of $\rho(\cdot, \cdot)$, (5.7) actually holds on the closed interval $[0, t_1]$.

Secondly, we show that (5.7) holds on the interval $[0, t_2]$ with

$$t_2 = \frac{1}{L} \ln \left\{ \frac{\delta}{2(1 + \rho(x_0, y_0))} + e^{Lt_1}(1 - L\epsilon_1) \right\},$$

where

$$\epsilon_1 \equiv \min \left\{ T, \frac{\delta}{3(1 + \rho(x_0, y_0))Le^{LT}} \right\}. \quad (5.12)$$

We already showed that (5.7) holds on $[0, t_1]$. It remains to show that, for $u(\cdot) \in \mathcal{U}_{ad}$ with $d(u(\cdot), \bar{u}(\cdot)) < \epsilon_1$, (5.7) still holds on $[t_1, t_2]$. For this purpose, for any $s \in [t_1, t_2]$, applying (5.7) with $t = t_1$ and (5.6), we see that

$$\begin{aligned} \rho(y_u(s), \bar{y}(s)) &\leq \rho(y_u(s), y_u(t_1)) + \rho(y_u(t_1), \bar{y}(t_1)) + \rho(\bar{y}(t_1), \bar{y}(s)) \\ &\leq 2(e^{Ls} - e^{Lt_1})(1 + \rho(x_0, y_0)) + 2L(1 + \rho(x_0, y_0))e^{Lt_1}d(u(\cdot), \bar{u}(\cdot)) \\ &< \delta. \end{aligned}$$

Then, similarly to (5.11), by Gronwall's inequality and (5.7) with $t = t_1$, we can get (5.7) on $[t_1, t_2]$. Furthermore, by the continuity of $\rho^2(\cdot, \cdot)$ and (5.6), one can easily get (5.7) on $[0, t_2]$.

Using the same technique as above, one can prove by induction that (5.7) holds on $[0, t_i]$ for $i = 2, 3, \dots$ with

$$t_i = \frac{1}{L} \ln \left\{ \frac{\delta}{2(1 + \rho(x_0, y_0))} + e^{Lt_{i-1}}(1 - L\epsilon_1) \right\}. \quad (5.13)$$

We claim that there exists an integer $I > 0$ being large enough such that $[0, T] \subset [0, t_I]$. Let us use the contradiction argument and assume that

$$t_i < T, \quad \forall i \in \{0\} \cup \mathbb{N}. \quad (5.14)$$

Then, there would exist a $\tilde{t} > 0$ and a subsequence of $\{t_i\}_{i=0}^{+\infty}$ (still denoted by $\{t_i\}_{i=0}^{+\infty}$), such that $\lim_{i \rightarrow +\infty} t_i = \tilde{t}$. Letting $i \rightarrow +\infty$ in (5.13) yields $\tilde{t} = \frac{1}{L} \ln \frac{\delta}{2(1 + \rho(x_0, y_0))Le^{\epsilon_1}}$. Recalling (5.12), one has $\tilde{t} > T$, which contradicts (5.14). Hence we get (5.7) on $[0, T]$. \square

We are now in a position to prove Proposition 5.1.

Proof of Proposition 5.1 First, we will prove (5.3). Multiplying both sides of (5.1) by $X_u(t)$, using (C1) and Lemma 4.1, we get

$$\begin{aligned} &\frac{1}{2}|X_u(t)|\frac{\partial}{\partial t}|X_u(t)| \\ &= \nabla_x f[t](X_u(t), X_u(t)) + \langle f(t, \bar{y}(t), u(t)) - f[t], X_u(t) \rangle \\ &\leq L|X_u(t)|^2 + |f(t, \bar{y}(t), u(t)) - f[t]||X_u(t)|, \end{aligned}$$

which leads to

$$|X_u(t)| \leq 2L \int_0^t |X_u(s)| ds + 2 \int_0^t |f(s, \bar{y}(s), u(s)) - f[s]| ds.$$

Applying Gronwall's inequality to the above inequality, we can get, via (C1), the first estimate in (5.3).

To prove the second estimate in (5.3), we take $Z = Y_u$ in (5.2), and get

$$\begin{aligned} \frac{1}{2}|Y_u(t)| \frac{\partial}{\partial t} |Y_u(t)| &= \nabla_x f[t](Y_u(t), Y_u(t)) + \nabla_x (f(t, \bar{y}(t), u(t)) - f[t])(Y_u(t), X_u(t)) \\ &\quad - \frac{1}{2}R(Y_u(t), X_u(t), f[t], X_u(t)) + \frac{1}{2}\nabla_x^2 f[t](\tilde{Y}_u(t), X_u(t), X_u(t)). \end{aligned}$$

Applying (C1), (C2), Lemma 4.1 and the first relation of (5.3), we can find a positive constant $C > 0$ such that

$$\frac{\partial}{\partial t} |Y_u(t)| \leq C(|Y_u(t)| + |\nabla_x f(t, \bar{y}(t), u(t)) - \nabla_x f[t]| |X_u(t)| + O(\epsilon^2)).$$

Integrating the above inequality on $[0, t]$, by Gronwall's inequality, and applying the first estimate in (5.3), the condition (C2) and Lemma 4.1, we obtain the second estimate in (5.3).

Next, we are going to prove (5.4).

By Lemma 5.1, there exists an $\epsilon_0 > 0$ such that, for all $u(\cdot) \in \mathcal{U}_{ad}$ with $d(u(\cdot), \bar{u}(\cdot)) = \epsilon < \epsilon_0$, we have $\rho(y_u(t), \bar{y}(t)) < i(\bar{y}(t))$ for all $t \in [0, T]$. Then, we can define $V_u(\cdot)$ as in (5.5). For $t \in [0, T]$, let $\tilde{V}_u(t)$ be a vector at $T_{\bar{y}(t)}M$ satisfying

$$\tilde{V}_u(t) \equiv \begin{cases} \frac{1}{\rho(\bar{y}(t), y_u(t))} V_u(t), & \text{if } |V_u(t)| \neq 0; \\ 0, & \text{if } |V_u(t)| = 0. \end{cases} \quad (5.15)$$

Then, we define a geodesic connecting $\bar{y}(t)$ and $y_u(t)$ as follows:

$$\beta(\theta; t) \equiv \exp_{\bar{y}(t)} \theta \tilde{V}_u(t), \quad \theta \in [0, \rho(\bar{y}(t), y_u(t))]. \quad (5.16)$$

The inequality (5.7) indicates that $\beta(\cdot; t)$ is the shortest geodesic connecting $\bar{y}(t)$ and $y_u(t)$, provided $d(\bar{u}(\cdot), u(\cdot)) = \epsilon$ is small enough. In particular, in the case that $|V_u(t)| \neq 0$, $\beta(\cdot; t)$ is parameterized by the arc length. For any $Z \in TM$, applying Lemma 2.2, Lemma 5.1

and Taylor's expansion, we get

$$\begin{aligned}
& \langle \nabla_{\dot{\bar{y}}(t)} V_u, Z \rangle \\
&= -\frac{1}{2} \nabla_2 \nabla_1 \rho^2(\bar{y}(t), y_u(t)) (Z, f(t, y_u(t), u(t))) + \frac{1}{2} \nabla_2 \nabla_1 \rho^2(\bar{y}(t), \bar{y}(t)) (Z, f(t, \bar{y}(t), u(t))) \\
&\quad - \frac{1}{2} \nabla_1^2 \rho^2(\bar{y}(t), y_u(t)) (Z, f[t]) + \frac{1}{2} \nabla_1^2 \rho^2(\bar{y}(t), \bar{y}(t)) (Z, f[t]) + \langle Z, f(t, \bar{y}(t), u(t)) - f[t] \rangle \\
&= -\frac{1}{2} \frac{\partial}{\partial \theta} |_{\theta=0} \nabla_2 \nabla_1 \rho^2(\bar{y}(t), \beta(\theta; t)) (Z, f(t, \beta(\theta; t), u(t))) \rho(\bar{y}(t), y_u(t)) \\
&\quad - \frac{1}{4} \frac{\partial^2}{\partial \theta^2} |_{\theta=0} \nabla_2 \nabla_1 \rho^2(\bar{y}(t), \beta(\theta; t)) (Z, f(t, \beta(\theta; t), u(t))) \rho^2(\bar{y}(t), y_u(t)) \\
&\quad - \frac{1}{2} \frac{\partial}{\partial \theta} |_{\theta=0} \nabla_1^2 \rho^2(\bar{y}(t), \beta(\theta; t)) (Z, f[t]) \rho(\bar{y}(t), y_u(t)) + \langle Z, f(t, \bar{y}(t), u(t)) - f[t] \rangle \\
&\quad - \frac{1}{4} \frac{\partial^2}{\partial \theta^2} |_{\theta=0} \nabla_1^2 \rho^2(\bar{y}(t), \beta(\theta; t)) (Z, f[t]) \rho^2(\bar{y}(t), y_u(t)) + o(\epsilon^2) \\
&= \langle Z, \nabla_{V_u(t)} f(t, \cdot, \bar{u}(t)) \rangle + \langle Z, f(t, \bar{y}(t), u(t)) - f[t] \rangle + \langle Z, \nabla_{V_u(t)} f(t, \cdot, u(t)) \\
&\quad - \nabla_{V_u(t)} f(t, \cdot, \bar{u}(t)) \rangle - \frac{1}{4} \nabla_2^3 \nabla_1 \rho^2(\bar{y}(t), \bar{y}(t)) (Z, f(t, \bar{y}(t), u(t)), V_u(t), V_u(t)) \\
&\quad - \frac{1}{4} \nabla_2^2 \nabla_1^2 \rho^2(\bar{y}(t), \bar{y}(t)) (Z, f[t], V_u(t), V_u(t)) \\
&\quad + \frac{1}{2} \langle Z, \nabla_{V_u(t)} \nabla_{\frac{\partial}{\partial \theta} \beta(\theta; t)} f(t, \cdot, u(t)) |_{\theta=0} \rangle \rho(\bar{y}(t), y_u(t)) + o(\epsilon^2).
\end{aligned}$$

By Lemma 2.3, we obtain

$$\begin{aligned}
& \langle \nabla_{\dot{\bar{y}}(t)} V_u, Z \rangle \\
&= \langle Z, \nabla_{V_u(t)} f(t, \cdot, \bar{u}(t)) \rangle + \langle Z, f(t, \bar{y}(t), u(t)) - f[t] \rangle + \langle Z, \nabla_{V_u(t)} f(t, \cdot, u(t)) \\
&\quad - \nabla_{V_u(t)} f(t, \cdot, \bar{u}(t)) \rangle - \frac{1}{2} R(f[t], V_u(t), Z, V_u(t)) \\
&\quad + \frac{1}{2} \langle Z, \nabla_{V_u(t)} \nabla_{\frac{\partial}{\partial \theta} \beta(\theta; t)} f(t, \cdot, u(t)) |_{\theta=0} \rangle \rho(\bar{y}(t), y_u(t)) \\
&\quad - \frac{1}{4} \nabla_2^3 \nabla_1 \rho^2(\bar{y}(t), \bar{y}(t)) (Z, f(t, \bar{y}(t), u(t)) - f[t], V_u(t), V_u(t)) + o(\epsilon^2).
\end{aligned} \tag{5.17}$$

We observe that, for any $v \in U$,

$$\langle Z, \nabla_{V_u(t)} \nabla_{\frac{\partial}{\partial \theta} \beta(\theta; t)} f(t, \cdot, v) |_{\theta=0} \rangle \rho(\bar{y}(t), y_u(t)) = \nabla_x^2 f(t, \bar{y}(t), v)(\tilde{Z}, V_u(t), V_u(t)). \tag{5.18}$$

In fact, recalling the definition of $\tilde{V}_u(t)$, (2.11), (2.10) and (5.16), we have

$$\begin{aligned}
& \text{The left hand side of (5.18)} \\
&= V_u(t) \langle \nabla_{\frac{\partial}{\partial \theta} \beta(\theta; t)} f(t, \cdot, v), Z \rangle \rho(y_u(t), \bar{y}(t)) - \langle \nabla_{\tilde{V}_u(t)} f(t, \cdot, v), \nabla_{V_u(t)} Z \rangle \rho(y_u(t), \bar{y}(t)) \\
&= V_u(t) \left(\nabla_x f(t, \cdot, v)(Z, \frac{\partial}{\partial \theta} \beta(\theta; t)) \right) \rho(y_u(t), \bar{y}(t)) - \nabla_x f(t, \bar{y}(t), v)(\nabla_{V_u(t)} Z, V_u(t)) \\
&= \left(\nabla_{V_u(t)} (\nabla_x f(t, \cdot, v)) \right) (Z, \frac{\partial}{\partial \theta} \beta(\theta; t) |_{\theta=0}) \rho(y_u(t), \bar{y}(t)) \\
&= \nabla_x^2 f(t, \bar{y}(t), v)(Z, V_u(t), V_u(t)),
\end{aligned}$$

which implies (5.18). Inserting (5.18) into (5.17), and subtracting (5.1) from (5.17), we get that

$$\left\{ \begin{aligned} & \langle \nabla_{\dot{\bar{y}}(t)} (V_u - X_u), Z \rangle = \langle Z, \nabla_{V_u(t) - X_u(t)} f(t, \cdot, \bar{u}(t)) \rangle - \frac{1}{2} R(f[t], V_u(t), Z, V_u(t)) \\ & \quad + \nabla_x \left(f(t, \bar{y}(t), u(t)) - f[t] \right) (\tilde{Z}, V_u(t)) + \frac{1}{2} \nabla_x^2 f[t] (\tilde{Z}, V_u(t), V_u(t)) \\ & \quad - \frac{1}{4} \nabla_2^3 \nabla_1 \rho^2(\bar{y}(t), \bar{y}(t)) (Z, f(t, \bar{y}(t), u(t)) - f[t], V_u(t), V_u(t)) \\ & \quad + \frac{1}{2} \nabla_x^2 (f(t, \bar{y}(t), u(t)) - f[t]) (\tilde{Z}, V_u(t), V_u(t)) + o(\epsilon^2), \quad t \in (0, T], \\ & V_u(0) - X_u(0) = 0. \end{aligned} \right. \tag{5.19}$$

Taking $Z = V_u(t) - X_u(t)$ in the above system, and using (C2), (C3), (5.3) and Lemma 2.2, we can find a positive constant $C > 0$ such that

$$\frac{\partial}{\partial t}|V_u(t) - X_u(t)|^2 \leq C|V_u(t) - X_u(t)|^2 + CI_{\{\bar{u} \neq u\}}(t)|V_u(t)|^2 + o(\epsilon^2),$$

where $I_{\{\bar{u} \neq u\}}$ is the indicator function of the set $\{\tau \in [0, T]; \bar{u}(\tau) \neq u(\tau)\}$, and the boundness of curvature tensor R along $\bar{y}(\cdot)$ is used. By integrating the above inequality over $[0, t]$, applying Gronwall's inequality of integral form and (5.3), we can get the first estimate in (5.4).

We subtract (5.2) from (5.19), and get

$$\left\{ \begin{array}{l} \langle \nabla_{\bar{y}(t)}(V_u - X_u - Y_u), Z \rangle = \langle Z, \nabla_{V_u(t) - X_u(t) - Y_u(t)} f(t, \cdot, \bar{u}(t)) \rangle \\ \quad + \nabla_x \left(f(t, \bar{y}(t), u(t)) - f[t] \right) (\tilde{Z}, V_u(t) - X_u(t)) \\ \quad - \frac{1}{2} R(Z, V_u(t) - X_u(t), f[t], X_u(t)) - \frac{1}{2} R(f[t], V_u(t) - X_u(t), Z, V_u(t)) \\ \quad + \frac{1}{2} \nabla_x^2 f[t] (\tilde{Z}, V_u(t) - X_u(t), V_u(t)) + \frac{1}{2} \nabla_x^2 f[t] (\tilde{Z}, X_u(t), V_u(t) - X_u(t)) \\ \quad - \frac{1}{4} \nabla_x^3 \nabla_1 \rho^2(\bar{y}(t), \bar{y}(t)) \left(Z, f(t, \bar{y}(t), u(t)) - f[t], V_u(t), V_u(t) \right) \\ \quad + \frac{1}{2} \nabla_x^2 (f(t, \bar{y}(t), u(t)) - f[t]) (\tilde{Z}, V_u(t), V_u(t)) + o(\epsilon^2), \quad t \in (0, T], \\ V_u(0) - X_u(0) - Y_u(0) = 0, \end{array} \right.$$

which implies

$$\begin{aligned} & |\nabla_{\bar{y}(t)}(V_u - X_u - Y_u)| \\ & \leq C|V_u(t) - X_u(t) - Y_u(t)| + CI_{\{u \neq \bar{u}\}}(t) \left(|V_u(t) - X_u(t)| + |V_u(t)|^2 \right) \\ & \quad + C|V_u(t) - X_u(t)|(|V_u(t)| + |X_u(t)|) + o(\epsilon^2), \end{aligned}$$

where we have used (C1), (C2), Lemma 4.1, the boundness of curvature tensor R along $\bar{y}(\cdot)$. Moreover, by applying (5.3), the first estimate in (5.4), Lemma 5.1 and (2.17), we can get

$$\begin{aligned} & \frac{\partial}{\partial t}|V_u(t) - X_u(t) - Y_u(t)|^2 \\ & \leq 2|\nabla_{\bar{y}(t)}(V_u - X_u - Y_u)||V_u(t) - X_u(t) - Y_u(t)| \\ & \leq C|V_u(t) - X_u(t) - Y_u(t)|^2 + C|V_u(t) - X_u(t) - Y_u(t)||V_u(t) - X_u(t)|(|V_u(t)| + |X_u(t)|) \\ & \quad + CI_{\{u \neq \bar{u}\}}(t) \left(|V_u(t) - X_u(t)| + |V_u(t)|^2 \right) |V_u(t) - X_u(t) - Y_u(t)| \\ & \quad + o(\epsilon^2)|V_u(t) - X_u(t) - Y_u(t)| \\ & \leq C|V_u(t) - X_u(t) - Y_u(t)|^2 + CI_{\{u \neq \bar{u}\}}(t)(|Y_u(t)|^2 + |V_u(t)|^4) + o(\epsilon^4). \end{aligned}$$

By integrating the above inequality over $[0, t]$ and applying Gronwall's inequality and (5.3), we can get the second estimate in (5.4). \square

5.2 Classical Variation

In this subsection, we assume that U is an open subset in \mathbb{R}^m . Similarly to Proposition 5.1, we have the following simple result.

Proposition 5.2 Assume that the conditions (C1), (C2) and (C3) hold. For any $\bar{u}(\cdot) \in L^2(0, T; U)$ and $v(\cdot) \in L^2(0, T; R^m)$, set

$$u^\epsilon(\cdot) = \bar{u}(\cdot) + \epsilon v(\cdot), \quad \epsilon \geq 0.$$

Denote by $y^\epsilon(\cdot)$ the solution to (1.1) with control $u^\epsilon(\cdot)$. In particular, we denote by $\bar{y}(\cdot)$ the solution to (1.1) with control $\bar{u}(\cdot)$. For any $\epsilon > 0$ being small enough, we define a vector field along $\bar{y}(\cdot)$ as follows:

$$V_\epsilon(t) = \exp_{\bar{y}(t)}^{-1} y^\epsilon(t), \quad t \in [0, T]. \quad (5.20)$$

Let $V(\cdot)$ and $Y(\cdot)$ be respectively the vector fields along $\bar{y}(\cdot)$ solving:

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} V(Z) = \nabla_x f[t](Z, V(t)) + \nabla_u f[t](Z, v(t)), & \text{a.e. } t \in (0, T], \forall Z \in T^*M, \\ V(0) = 0, \end{cases} \quad (5.21)$$

and

$$\begin{cases} \nabla_{\dot{\bar{y}}(t)} Y(Z) = \nabla_x f[t](Z, Y(t)) + \nabla_x \nabla_u f[t](Z, v(t), Y(t)) - \frac{1}{2} R(\tilde{Z}, V(t), \dot{\bar{y}}(t), V(t)) \\ \quad + \frac{1}{2} \nabla_x^2 f[t](Z, V(t), V(t)) + \frac{1}{2} \nabla_u^2 f[t](Z, v(t), v(t)), & \text{a.e. } t \in (0, T], \forall Z \in T^*M, \\ Y(0) = 0. \end{cases} \quad (5.22)$$

Then, we have

$$V_\epsilon(t) = \epsilon V(t) + \epsilon^2 Y(t) + o(\epsilon^2), \quad \forall t \in [0, T]. \quad (5.23)$$

Proof Similarly to Lemma 5.1, when $\epsilon > 0$ is small enough, $\rho(y^\epsilon(t), \bar{y}(t)) = O(\epsilon)$ for any $t \in [0, T]$. Thus, one can define the vector field (5.20).

Following a similar argument as that in the proof of (5.17), we can obtain, for any $Z \in T_{\bar{y}(t)} M$,

$$\begin{aligned} & \langle \nabla_{\dot{\bar{y}}(t)} V_\epsilon, Z \rangle \\ &= \nabla_x f[t](Z, V_\epsilon(t)) + \nabla_u f[t](Z, v(t))\epsilon + \nabla_x \nabla_u f[t](Z, v(t), V_\epsilon(t))\epsilon - \frac{1}{2} R(Z, V_\epsilon(t), \dot{\bar{y}}(t), V_\epsilon(t)) \\ & \quad + \frac{1}{2} \nabla_x^2 f[t](Z, V_\epsilon(t), V_\epsilon(t)) + \frac{1}{2} \nabla_u^2 f[t](Z, v(t), v(t))\epsilon^2 + o(\epsilon^2). \end{aligned}$$

Dividing the above identity by ϵ , we obtain that

$$\lim_{\epsilon \rightarrow 0^+} \frac{V_\epsilon(t)}{\epsilon} = V(t) \quad (5.24)$$

uniformly in $t \in [0, T]$, where $V(\cdot)$ is the solution to (5.21).

Similarly, one can check that

$$\lim_{\epsilon \rightarrow 0^+} \frac{V_\epsilon(t) - \epsilon V(t)}{\epsilon^2} = Y(t)$$

uniformly in $t \in [0, T]$, where $Y(\cdot)$ is a vector field along $\bar{y}(\cdot)$ solving (5.22). Thus, the proof is concluded. \square

6 Proof of the Main Results

This section is addressed to proving our main results, i.e., Theorems 3.1–3.4.

6.1 Proof of Theorem 3.1

The proof of Pontryagin-type maximum principle (3.11) is given in [1, p. 183, Theorem 12.13]. This subsection aims at proving the second order necessary condition. We divide the proof into 3 steps.

Step 1 In this step, we will give the second order expansion of $J(\cdot)$ around the optimal control $\bar{u}(\cdot)$. Fix any $u(\cdot) \in \mathcal{U}_{ad}$ and $\epsilon > 0$. Define

$$u^\epsilon(t) = \bar{u}(t)I_{[0,T] \setminus E_\epsilon}(t) + u(t)I_{E_\epsilon}(t), \quad (6.1)$$

where $E_\epsilon \subset [0, T]$ is a measurable subset of $[0, T]$ with $|E_\epsilon| = \epsilon$. Let $y^\epsilon(\cdot)$ be the solution to (1.1) corresponding to the control $u^\epsilon(\cdot)$. Denote by $X^\epsilon(\cdot)$ and $Y^\epsilon(\cdot)$ the solutions to (5.1) and (5.2) with $u(\cdot) = u^\epsilon(\cdot)$, respectively. By Proposition 5.1, there exists an $\epsilon_0 > 0$ such that, for all $\epsilon < \epsilon_0$, we can define

$$V^\epsilon(t) \equiv \exp_{\bar{y}(t)}^{-1} y^\epsilon(t), \quad \forall t \in [0, T],$$

and

$$X^\epsilon(t) = O(\epsilon), \quad Y^\epsilon(t) = O(\epsilon^2), \quad V^\epsilon(t) = X^\epsilon(t) + Y^\epsilon(t) + o(\epsilon^2), \quad \forall t \in [0, T]. \quad (6.2)$$

Similar to (5.15) and (5.16), for $t \in [0, T]$, let $\tilde{V}^\epsilon(t)$ be a vector at $\bar{y}(t)$ satisfying

$$\tilde{V}^\epsilon(t) \equiv \begin{cases} \frac{1}{\rho(\bar{y}(t), y^\epsilon(t))} V^\epsilon(t), & \text{if } |V^\epsilon(t)| \neq 0; \\ 0, & \text{if } |V^\epsilon(t)| = 0, \end{cases} \quad (6.3)$$

and let $\gamma(\cdot; t)$ be the shortest geodesic connecting $\bar{y}(t)$ and $y^\epsilon(t)$, given by

$$\gamma(\theta; t) = \exp_{\bar{y}(t)}(\theta \tilde{V}^\epsilon(t)), \quad \theta \in [0, \rho(\bar{y}(t), y^\epsilon(t))]. \quad (6.4)$$

In particular, in the case that $|V^\epsilon(t)| \neq 0$, $\gamma(\cdot; t)$ is parameterized by the arc length. Recalling (3.6), and applying (2.4), Taylor's Theorem, Lemma 5.1, (6.3) and (6.4), we

obtain that

$$\begin{aligned}
& J(u^\epsilon(\cdot)) - J(\bar{u}(\cdot)) \\
&= \int_0^T \left(f^0(t, y^\epsilon(t), u^\epsilon(t)) - f^0(t, \bar{y}(t), u^\epsilon(t)) + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right) dt \\
&= \int_0^T \left(f^0(t, \gamma(\rho(\bar{y}(t), y^\epsilon(t)); t), u^\epsilon(t)) - f^0(t, \gamma(0; t), u^\epsilon(t)) \right. \\
&\quad \left. + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right) dt \\
&= \int_0^T \left(\frac{\partial}{\partial s} \Big|_0 f^0(t, \gamma(s; t), u^\epsilon(t)) \rho(\bar{y}(t), y^\epsilon(t)) + \frac{1}{2} \frac{\partial^2}{\partial s^2} \Big|_0 f^0(t, \gamma(s; t), u^\epsilon(t)) \rho^2(\bar{y}(t), y^\epsilon(t)) \right. \\
&\quad \left. + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] + o(\rho^2(\bar{y}(t), y^\epsilon(t))) \right) dt \\
&= \int_0^T \left(\langle \nabla_x f^0(t, \bar{y}(t), u^\epsilon(t)), \tilde{V}^\epsilon(t) \rangle \rho(\bar{y}(t), y^\epsilon(t)) \right. \\
&\quad \left. + \frac{1}{2} \nabla_x^2 f^0(t, \bar{y}(t), u^\epsilon(t)) (\tilde{V}^\epsilon(t), \tilde{V}^\epsilon(t)) \rho^2(\bar{y}(t), y^\epsilon(t)) + f^0(t, \bar{y}(t), u^\epsilon(t)) \right. \\
&\quad \left. - f^0[t] + o(\rho^2(\bar{y}(t), y^\epsilon(t))) \right) dt \\
&= \int_0^T \left(\langle \nabla_x f^0[t], V^\epsilon(t) \rangle + \frac{1}{2} \nabla_x^2 f^0[t] (V^\epsilon(t), V^\epsilon(t)) + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right. \\
&\quad \left. + \frac{1}{2} \nabla_x^2 \left(f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0(t, \bar{y}(t), \bar{u}(t)) \right) (V^\epsilon(t), V^\epsilon(t)) \right. \\
&\quad \left. + \langle \nabla_x f^0(t, \bar{y}(t), u^\epsilon(t)) - \nabla_x f^0[t], V^\epsilon(t) \rangle + o(\rho^2(\bar{y}(t), y^\epsilon(t))) \right) dt.
\end{aligned}$$

Recalling (5.7) and (6.2), we further get

$$J(u^\epsilon(\cdot)) - J(\bar{u}(\cdot)) = I_1 + I_2 + o(\epsilon^2), \quad (6.5)$$

where

$$I_1 = \int_0^T \left(\langle \nabla_x f^0[t], X^\epsilon(t) \rangle + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right) dt, \quad (6.6)$$

$$\begin{aligned}
I_2 = \int_0^T & \left(\langle \nabla_x f^0[t], Y^\epsilon(t) \rangle + \langle \nabla_x f^0(t, \bar{y}(t), u^\epsilon(t)) - \nabla_x f^0[t], X^\epsilon(t) \rangle \right. \\
& \left. + \frac{1}{2} \nabla_x^2 f^0[t] (X^\epsilon(t), X^\epsilon(t)) \right) dt.
\end{aligned} \quad (6.7)$$

In what follows, we will rewrite I_1 and I_2 by the dual variables and the perturbed control $u^\epsilon(\cdot)$. Recalling the first order variational equation (5.1), the first order dual equation (3.7) and (3.4) with $\nu = -1$ and $\psi_1 = 0$, we have, via integration by parts over $[0, T]$,

$$\begin{aligned}
I_1 &= \int_0^T \left(\nabla_{\dot{y}(t)} \psi(X^\epsilon(t)) + \nabla_{X^\epsilon(t)} f(t, \cdot, \bar{u}(t))(\psi(t)) + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right) dt \\
&= \int_0^T \left(\psi(t) \left(-\nabla_{\dot{y}(t)} X^\epsilon + \nabla_{X^\epsilon(t)} f(t, \cdot, \bar{u}(t)) \right) + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right) dt \\
&= \int_0^T \left(\psi(t) \left(f[t] - f(t, \bar{y}(t), u^\epsilon(t)) \right) + f^0(t, \bar{y}(t), u^\epsilon(t)) - f^0[t] \right) dt \\
&= \int_0^T \left(H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - H(t, \bar{y}(t), \psi(t), u^\epsilon(t)) \right) dt.
\end{aligned} \quad (6.8)$$

Using (3.7) with $\nu = -1$ and $\psi_1 = 0$, the second order dual equation (3.8), the second order variational equation (5.2) and integration by parts over $[0, T]$, via (3.13) and (3.16),

we can get

$$\begin{aligned}
I_2 &= \int_0^T \left(\nabla_{\dot{\bar{y}}(t)} \psi(Y^\epsilon(t)) + \psi(t)(\nabla_{Y^\epsilon(t)} f(t, \cdot, \bar{u}(t))) + \langle \nabla_x f^0(t, \bar{y}(t), u^\epsilon(t)) \right. \\
&\quad \left. - \nabla_x f^0[t], X^\epsilon(t) \rangle + \frac{1}{2} \nabla_x^2 f^0[t](X^\epsilon(t), X^\epsilon(t)) \right) dt \\
&= \int_0^T \left(\psi(t) \left(-\nabla_{\dot{\bar{y}}(t)} Y^\epsilon + \nabla_{Y^\epsilon(t)} f(t, \cdot, \bar{u}(t)) \right) + \langle \nabla_x f^0(t, \bar{y}(t), u^\epsilon(t)) \right. \\
&\quad \left. - \nabla_x f^0[t], X^\epsilon(t) \rangle + \frac{1}{2} \nabla_x^2 f^0[t](X^\epsilon(t), X^\epsilon(t)) \right) dt \\
&= \int_0^T \left(\frac{1}{2} R(\tilde{\psi}(t), X^\epsilon(t), f[t], X^\epsilon(t)) - \langle \nabla_{X^\epsilon(t)} (f(t, \cdot, u^\epsilon(t)) - f(t, \cdot, \bar{u}(t))), \tilde{\psi}(t) \rangle \right. \\
&\quad \left. - \frac{1}{2} \nabla_x^2 f[t](\psi(t), X^\epsilon(t), X^\epsilon(t)) + \langle \nabla_x f^0(t, \bar{y}(t), u^\epsilon(t)) - \nabla_x f^0[t], X^\epsilon(t) \rangle \right. \\
&\quad \left. + \frac{1}{2} \nabla_x^2 f^0[t](X^\epsilon(t), X^\epsilon(t)) \right) dt \\
&= I_3 + I_4,
\end{aligned} \tag{6.9}$$

where $\tilde{\psi}$ is the dual vector of ψ ,

$$I_3 = \frac{1}{2} \int_0^T \left(R(\tilde{\psi}(t), X^\epsilon(t), f[t], X^\epsilon(t)) - \nabla_x^2 H(t, \bar{y}(t), \psi(t), \bar{u}(t))(X^\epsilon(t), X^\epsilon(t)) \right) dt, \tag{6.10}$$

and

$$I_4 = \int_0^T \langle \nabla_x H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - \nabla_x H(t, \bar{y}(t), \psi(t), u^\epsilon(t)), X^\epsilon(t) \rangle dt. \tag{6.11}$$

Next, we shall re-write I_3 and I_4 in terms of local coordinates along $\bar{y}(\cdot)$. For this purpose, let $\{e_1, e_2, \dots, e_n\}$ be an orthonormal basis at $\bar{y}(0) = y_0$, and $\{d_1, d_2, \dots, d_n\}$ be the dual basis of $\{e_1, e_2, \dots, e_n\}$. For $t \in [0, T]$, we define

$$e_i(t) \equiv L_{y_0 \bar{y}(t)}^{\bar{y}(\cdot)} e_i, \quad d_i(t) \equiv L_{y_0 \bar{y}(t)}^{\bar{y}(\cdot)} d_i, \quad i = 1, 2, \dots, n. \tag{6.12}$$

By (2.6) and (2.7), we have

$$\langle e_i(t), e_j(t) \rangle = d_i(t)(e_j(t)) = \delta_i^j, \quad i, j = 1, 2, \dots, n, \quad t \in [0, T]. \tag{6.13}$$

Set

$$X^\epsilon(t) = \sum_{i=1}^n X_i^\epsilon(t) e_i(t), \quad \vec{X}^\epsilon(t) = (X_1^\epsilon(t), \dots, X_n^\epsilon(t))^\top. \tag{6.14}$$

Inserting the above equalities into (6.10), we get

$$I_3 = \frac{1}{2} \text{tr} \int_0^T (M(t) - \vec{H}(t)) \vec{X}^\epsilon(t) \vec{X}^{\epsilon^\top}(t) dt, \tag{6.15}$$

where

$$M(t) = (M_{ik}(t)) \equiv \left(R(\tilde{\psi}(t), e_i(t), f[t], e_k(t)) \right), \quad t \in [0, T], \tag{6.16}$$

$$\vec{H}(t) = (H_{ik}(t)) \equiv \left(\nabla_x^2 H(t, \bar{y}(t), \psi(t), \bar{u}(t))(e_i(t), e_k(t)) \right), \quad t \in [0, T]. \tag{6.17}$$

Let us derive the differential equation satisfied by the matrix $\vec{X}^\epsilon(t) \vec{X}^{\epsilon^\top}(t)$. Inserting (6.14) into (5.1), via the parallel translation of $e_i(\cdot)$ ($i = 1, \dots, n$) along $\bar{y}(\cdot)$, we can get

$$\begin{cases} \dot{\vec{X}}^\epsilon(t) = F(t) \vec{X}^\epsilon(t) + F_1(t, u^\epsilon(t)), & t \in (0, T], \\ \vec{X}^\epsilon(0) = 0, \end{cases} \tag{6.18}$$

where

$$F(t) = \left(\left\langle \nabla_{e_j(t)} f(t, \cdot, \bar{u}(t)), e_i(t) \right\rangle \right) = \nabla_x f[t](d_i(t), e_j(t)) \equiv (F_{ij}(t)), \quad (6.19)$$

$$F_1(t, u) = \begin{pmatrix} \langle f(t, \bar{y}(t), u) - f[t], e_1(t) \rangle \\ \vdots \\ \langle f(t, \bar{y}(t), u) - f[t], e_n(t) \rangle \end{pmatrix} \equiv (F_1^1(t, u), \dots, F_1^n(t, u))^\top. \quad (6.20)$$

Hence, we get the equation for $\vec{X}^\epsilon(t) \vec{X}^{\epsilon\top}(t)$ as follows:

$$\begin{cases} \frac{d}{dt}(\vec{X}^\epsilon \vec{X}^{\epsilon\top})(t) = F(t)(\vec{X}^\epsilon \vec{X}^{\epsilon\top})(t) + (\vec{X}^\epsilon \vec{X}^{\epsilon\top})(t)F(t)^\top + F_1(t, u^\epsilon(t))\vec{X}^\epsilon(t)^\top \\ \quad + \vec{X}^\epsilon(t)F_1(t, u^\epsilon(t))^\top, \quad t \in (0, T], \\ (\vec{X}^\epsilon \vec{X}^{\epsilon\top})(0) = 0. \end{cases} \quad (6.21)$$

We now derive the local form of the second order dual equation (3.8). Recalling (6.13), we can express the solution w to (3.8) by

$$w(t) = \sum_{i,j=1}^n w_{ij}(t) d_i(t) \otimes d_j(t), \quad (6.22)$$

where $w_{ij}(t) \equiv w(t)(e_i(t), e_j(t))$ for any $t \in [0, T]$. The transpose of $w(t)$ can be rewritten as

$$w^\top(t) = \sum_{i,j=1}^n w_{ji}(t) d_i(t) \otimes d_j(t), \quad \forall t \in [0, T]. \quad (6.23)$$

Recalling (6.19), we have

$$\nabla_x f[t] = \sum_{i,j=1}^n F_{ij}(t) e_i(t) \otimes d_j(t). \quad (6.24)$$

Thanks to (6.16) and (6.17), we have

$$R(\tilde{\psi}(t), \cdot, f[t], \cdot) = \sum_{i,j=1}^n M_{ij}(t) d_i(t) \otimes d_j(t), \quad (6.25)$$

$$\nabla_x^2 H(t, \bar{y}(t), \psi(t), \bar{u}(t)) = \sum_{i,j=1}^n H_{ij}(t) d_i(t) \otimes d_j(t). \quad (6.26)$$

Set

$$W(t) = (w_{ij}(t)), \quad \forall t \in [0, T]. \quad (6.27)$$

Recalling (6.12) and (6.13), we insert (6.22), (6.24), (6.25) and (6.26) into (3.8), and get

$$\begin{cases} \dot{W}(t) + F(t)^\top W(t) + W(t)F(t) - M(t) + \vec{H}(t) = 0, \quad t \in [0, T], \\ W(T) = 0. \end{cases} \quad (6.28)$$

Applying (6.21), (6.28) to (6.15), via integration by parts over $[0, T]$, we get

$$\begin{aligned}
I_3 &= \frac{1}{2} \text{tr} \int_0^T \left(\dot{W}(t) + F(t)^\top W(t) + W(t)F(t) \right) \vec{X}^\epsilon(t) \vec{X}^{\epsilon^\top}(t) dt \\
&= \frac{1}{2} \text{tr} \int_0^T \left\{ W(t) \left(-\frac{d}{dt} (\vec{X}^\epsilon \vec{X}^{\epsilon^\top})(t) + F(t) (\vec{X}^\epsilon \vec{X}^{\epsilon^\top})(t) \right) + F(t)^\top W(t) (\vec{X}^\epsilon \vec{X}^{\epsilon^\top})(t) \right\} dt \\
&= \frac{1}{2} \text{tr} \int_0^T \left\{ -W(t) (\vec{X}^\epsilon \vec{X}^{\epsilon^\top})(t) F(t)^\top - W(t) F_1(t, u^\epsilon(t)) \vec{X}^{\epsilon^\top}(t) - W(t) \vec{X}^\epsilon(t) F_1(t, u^\epsilon(t))^\top \right. \\
&\quad \left. + F(t)^\top W(t) \vec{X}^\epsilon(t) \vec{X}^{\epsilon^\top}(t) \right\} dt \\
&= -\frac{1}{2} \text{tr} \int_0^T (W(t) + W^\top(t)) F_1(t, u^\epsilon(t)) \vec{X}^{\epsilon^\top}(t) dt \\
&= -\frac{1}{2} \int_0^T \left((W(t) + W^\top(t)) F_1(t, u^\epsilon(t)) \cdot \vec{X}^\epsilon(t) \right) dt,
\end{aligned} \tag{6.29}$$

where we have used the properties: $\text{tr}(AB) = \text{tr}(BA)$ and $\text{tr}A = \text{tr}(A^\top)$ for any $n \times n$ matrixes A and B .

In order to express I_4 in the local form, for any $u \in U$ and $i = 1, \dots, n$, set

$$\delta \partial_i H(t, u) = \left(\nabla_x H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - \nabla_x H(t, \bar{y}(t), \psi(t), u) \right) (e_i(t)), \tag{6.30}$$

$$\delta \partial H(t, u) = \left(\delta \partial_1 H(t, u), \dots, \delta \partial_n H(t, u) \right)^\top. \tag{6.31}$$

Recalling (6.14), we have

$$I_4 = \int_0^T \delta \partial H(t, u^\epsilon(t)) \cdot \vec{X}^\epsilon(t) dt. \tag{6.32}$$

Now, we re-write \vec{X}^ϵ in terms of $u^\epsilon(\cdot)$. Assume that $\tilde{\Phi}(\cdot) = (\phi_{ij}(\cdot))$ is an $\mathbb{R}^{n \times n}$ -valued function satisfying

$$\begin{cases} \dot{\tilde{\Phi}}(t) = F(t) \tilde{\Phi}(t), & t \in (0, T], \\ \tilde{\Phi}(0) = I, \end{cases} \tag{6.33}$$

where $I \in \mathbb{R}^{n \times n}$ is the $n \times n$ identity matrix. Recalling (6.18), we have

$$\vec{X}^\epsilon(t) = \int_0^t \tilde{\Phi}(t) \tilde{\Phi}^{-1}(s) F_1(s, u^\epsilon(s)) ds, \tag{6.34}$$

where $\tilde{\Phi}^{-1}(\cdot) = (\psi_{ij}(\cdot))$ is the inverse of $\tilde{\Phi}(\cdot)$.

Inserting (6.29) and (6.32) into (6.9) via (6.34), and recalling (6.1), (6.8), (6.20), (6.30) and (6.31), we can rewrite (6.5) by

$$\begin{aligned}
&J(u^\epsilon(\cdot)) - J(\bar{u}(\cdot)) \\
&= \int_{E_\epsilon} \left(H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - H(t, \bar{y}(t), \psi(t), u(t)) \right) dt + \int_{E_\epsilon} \left\{ -\frac{1}{2} (W(t) \right. \\
&\quad \left. + W^\top(t)) F_1(t, u(t)) + \delta \partial H(t, u(t)) \right\} \cdot \tilde{\Phi}(t) \int_0^t \tilde{\Phi}^{-1}(s) F_1(s, u^\epsilon(s)) ds dt + o(\epsilon^2).
\end{aligned} \tag{6.35}$$

Step 2 In order to get the integral form of second order necessary condition, we apply [25, Corollary 3.8, p. 144] to choose suitable measurable set E_ϵ in (6.35). Let $u(\cdot)$ be such

that $u(t) \in \tilde{U}(t)$, a.e. $t \in [0, T]$. Let $E_\epsilon \subset [0, T]$ with $|E_\epsilon| = \epsilon$ be such that

$$\begin{aligned} \int_{[0,t] \cap E_\epsilon} \tilde{\Phi}^{-1}(s) F_1(s, u(s)) ds &= \epsilon \int_0^t \tilde{\Phi}^{-1}(s) F_1(s, u(s)) ds + \eta_1(t), \\ \int_{[0,t] \cap E_\epsilon} \left\{ -\frac{1}{2}(W(\tau) + W^\top(\tau)) F_1(\tau, u(\tau)) \right. \\ &\quad \left. + \delta \partial H(\tau, u(\tau)) \right\} \cdot \tilde{\Phi}(\tau) \int_0^\tau \tilde{\Phi}^{-1}(s) F_1(s, u(s)) ds d\tau \\ &= \epsilon \int_0^t \left\{ -\frac{1}{2}(W(\tau) + W^\top(\tau)) F_1(\tau, u(\tau)) \right. \\ &\quad \left. + \delta \partial H(\tau, u(\tau)) \right\} \cdot \tilde{\Phi}(\tau) \int_0^\tau \tilde{\Phi}^{-1}(s) F_1(s, u(s)) ds d\tau + \eta_2(t), \end{aligned}$$

where

$$|\eta_i(t)| \leq \epsilon^3, \quad \forall t \in [0, T], i = 1, 2.$$

Then, dividing (6.35) by ϵ^2 and taking the limit as $\epsilon \rightarrow 0$, one can get, via the optimality of $\bar{u}(\cdot)$,

$$0 \leq \int_0^T \left\{ -\frac{1}{2}(W(t) + W^\top(t)) F_1(t, u(t)) + \delta \partial H(t, u(t)) \right\} \cdot \tilde{\Phi}(t) \int_0^t \tilde{\Phi}^{-1}(s) F_1(s, u(s)) ds dt. \quad (6.36)$$

Set

$$\Phi(t) = \sum_{i,j=1}^n \phi_{ij}(t) e_i(t) \otimes e_j(t), \quad \Phi_1(t) = \sum_{i,j=1}^n \psi_{ij}(t) d_i(t) \otimes d_j(t), \quad t \in [0, T]. \quad (6.37)$$

Recalling (2.16), (6.24) and (6.33), one can check, via (6.12) and (2.7), that $\Phi(\cdot)$ and $\Phi_1(\cdot)$ solve (3.9) and (3.10) respectively.

Recalling (6.20), (6.22), (6.23), (6.27), (6.30), (6.31) and (6.37), one can get (3.12) from (6.36), via (6.12) and (6.13).

Step 3 In this step, we shall get the pointwise form of the second order necessary condition. Assume that U is a Polish space. Recalling (6.36), define

$$\begin{aligned} G_1(t, u) &= \tilde{\Phi}(t)^\top \left\{ -\frac{1}{2}(W(t) + W^\top(t)) F_1(t, u) + \delta \partial H(t, u) \right\}, \\ G_2(t, u) &= \tilde{\Phi}^{-1}(t) F_1(t, u), \quad (t, u) \in [0, T] \times U. \end{aligned} \quad (6.38)$$

Then, by (6.36), for any $u(\cdot) \in \mathcal{U}_{ad}$ and $u(t) \in \tilde{U}(t)$, a.e. $t \in [0, T]$, one has

$$0 \leq \int_0^T G_1(t, u(t)) \cdot \int_0^t G_2(s, u(s)) ds dt. \quad (6.39)$$

Applying the same argument as that in the proof of [26, Theorem 4.3], we can get, from the above inequality and the conditions (C1) and (C2),

$$G_1(t, u(t)) \cdot G_2(t, u(t)) \geq 0 \quad a.e. t \in [0, T].$$

Following the same argument as that in the proof of [26, Theorem 4.3] again, we can get, from the above inequality and the assumption that U is a Polish space,

$$G_1(t, v) \cdot G_2(t, v) \geq 0 \quad \forall v \in \tilde{U}(t),$$

for a.e. $t \in [0, T]$. Recalling (6.38), the above inequality implies

$$0 \leq \left\{ -\frac{1}{2}(W(t) + W^\top(t))F_1(t, v) + \delta\partial H(t, v) \right\} \cdot F_1(t, v), \quad \forall v \in \tilde{U}(t), \text{ a.e. } t \in [0, T]. \quad (6.40)$$

Combining the above inequality with (6.13), (6.20), (6.22), (6.23), (6.30) and (6.31), we can easily obtain (3.14). \square

6.2 Proof of Theorem 3.2

For any $v(\cdot) \in \mathcal{U}_{ad}$ with $d(\bar{u}(\cdot), v(\cdot)) = \epsilon$ being small enough, from (6.35), we have,

$$\begin{aligned} & J(v(\cdot)) - J(\bar{u}(\cdot)) \\ &= \int_0^T \left(H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - H(t, \bar{y}(t), \psi(t), v(t)) \right) dt + \int_0^T \left\{ -\frac{1}{2}(W(t) + W^\top(t)) \right. \\ & \quad \left. F_1(t, v(t)) + \delta\partial H(t, v(t)) \right\} \cdot \tilde{\Phi}(t) \int_0^t \tilde{\Phi}^{-1}(s) F_1(s, v(s)) ds dt + o(\epsilon^2). \end{aligned}$$

Recalling (6.20), (6.22), (6.23), (6.27), (6.30), (6.31) and (6.37), one can get, via (6.12) and (6.13),

$$\begin{aligned} & J(v(\cdot)) - J(\bar{u}(\cdot)) \\ &= \int_0^T \left(H(t, \bar{y}(t), \psi(t), \bar{u}(t)) - H(t, \bar{y}(t), \psi(t), v(t)) \right) dt - \int_0^T \int_0^t \left\{ \frac{1}{2}(w(t) + w^\top(t)) \right. \\ & \quad \left(f(t, \bar{y}(t), v(t)) - f[t], \mathcal{E}^{21} \left(\Phi(t) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} \Phi_1(s) \right) \otimes \right. \\ & \quad \left. L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} (f(s, \bar{y}(s), v(s)) - f[s]) \right) \Big) + \left\langle \nabla_x H(t, \bar{y}(t), \psi(t), v(t)) \right. \\ & \quad \left. - \nabla_x H(t, \bar{y}(t), \psi(t), \bar{u}(t)), \mathcal{E}^{21} \left(\Phi(t) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} \Phi_1(s) \right) \otimes L_{\bar{y}(s)\bar{y}(t)}^{\bar{y}(\cdot)} (f(s, \bar{y}(s), v(s)) \right. \\ & \quad \left. \left. - f(s, \bar{y}(s), \bar{u}(s))) \right) \right\rangle \Big\} ds dt + o(\epsilon^2). \end{aligned}$$

Hence, according to (3.11) and (3.15), it is easy to see that, there exists an $\epsilon_1 > 0$ such that $J(u^\epsilon(\cdot)) - J(\bar{u}(\cdot)) \geq 0$, for all $u^\epsilon(\cdot) \in \mathcal{U}_{ad}$ with $d(u^\epsilon(\cdot), \bar{u}(\cdot)) = \epsilon \leq \epsilon_1$. This completes the proof of Theorem 3.2. \square

6.3 Proofs of Theorems 3.3 and 3.4

In this subsection, we shall prove Theorems 3.3 and 3.4.

To begin with, let us recall the following known result ([1, Theorem 20.3, p. 297]):

Lemma 6.1 *Let $F : \mathcal{U} \rightarrow \mathcal{M}$ be a continuous mapping having smooth restrictions to finite-dimensional submanifolds of \mathcal{U} , where \mathcal{U} is an open subset of a Banach space, and \mathcal{M} is an n -dimensional, smooth differential manifold. Let $\hat{u} \in \mathcal{U}$ be a corank one critical point of F , i.e. the codimension of $\text{Im} D_{\hat{u}} F$ (the image of the differential of F at \hat{u}) is equal to 1. Let $\lambda \in (\text{Im} D_{\hat{u}} F)^\perp \equiv \{\eta \in T_{F(\hat{u})}^* \mathcal{M}; \quad \eta(X) = 0, \forall X \in \text{Im} D_{\hat{u}} F \subset T_{F(\hat{u})} \mathcal{M}\}$, $\lambda \neq 0$. If the quadratic form $\lambda \text{Hess}_{\hat{u}} F : \ker D_{\hat{u}} F \times \ker D_{\hat{u}} F \rightarrow \mathbb{R}$ is sign-indefinite, which is defined by*

$$\lambda \text{Hess}_{\hat{u}} F(v, v) \equiv \lambda \left(\frac{d^2}{d\epsilon^2} \Big|_0 F(\varphi(\epsilon)) \right), \quad \forall v \in \text{Ker} D_{\hat{u}} F \subset T_{\hat{u}} \mathcal{U}, \quad (6.41)$$

with $\varphi : [0, \epsilon_0) \rightarrow \mathcal{U}$ ($\epsilon_0 > 0$) satisfying $\varphi(0) = \hat{u}$ and $\frac{d}{d\epsilon}\big|_0 \varphi(\epsilon) = v$, then F is locally open at \hat{u} , i.e. $F(\hat{u}) \in \text{int}F(O_{\hat{u}})$ for any neighborhood $O_{\hat{u}} \subset \mathcal{U}$ of \hat{u} .

We are now in a position to prove Theorem 3.3.

Proof of Theorem 3.3 Similarly to [1], we introduce an extended system associated to problem (1.1) and (1.2) as follows:

$$\begin{cases} \frac{d}{dt} \begin{pmatrix} y_0(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} f^0(t, y(t), u(t)) \\ f(t, y(t), u(t)) \end{pmatrix}, & u(t) \in U, \text{ a.e. } t \in (0, T), \\ \begin{pmatrix} y_0(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 0 \\ y_0 \end{pmatrix}. \end{cases} \quad (6.42)$$

Define the endpoint mapping $E : L^2(0, T; U) \rightarrow \mathbb{R} \times M$ for the system (6.42) by

$$E(u(\cdot)) \equiv \begin{pmatrix} J(u(\cdot)) \\ y(T; u(\cdot)) \end{pmatrix}, \quad (6.43)$$

where $y(\cdot; u(\cdot))$ is the solution to (1.1) associated to the control $u(\cdot)$. Then, we define the attainable set of (6.42) at time T by

$$\mathcal{A} \equiv \{E(u(\cdot)); \ u(\cdot) \in L^2(0, T; U)\}.$$

Since $\bar{u}(\cdot)$ is optimal for **Problem II**, we have ([1, Section 12.4, p. 179])

$$E(\bar{u}(\cdot)) \in \partial\mathcal{A}. \quad (6.44)$$

For any $v(\cdot) \in L^2(0, T; \mathbb{R}^m)$, let $y^\epsilon(\cdot)$ be the solution to (1.1) corresponding to the control $\bar{u}(\cdot) + \epsilon v(\cdot)$ with small $\epsilon \geq 0$.

Firstly, we claim that

$$\frac{\partial}{\partial \epsilon}\bigg|_0 y^\epsilon(t) = V(t), \quad \frac{\partial^2}{\partial \epsilon^2}\bigg|_0 y^\epsilon(t) = 2Y(t), \quad t \in [0, T], \quad (6.45)$$

where $V(\cdot)$ and $Y(\cdot)$ are the solutions to (5.21) and (5.22), respectively.

In fact, for any $t \in [0, T]$, set

$$\tilde{V}_\epsilon(t) \equiv \begin{cases} \frac{V_\epsilon(t)}{\rho(\bar{y}(t), y^\epsilon(t))}, & \text{if } \rho(\bar{y}(t), y^\epsilon(t)) > 0, \\ 0, & \text{if } \rho(\bar{y}(t), y^\epsilon(t)) = 0, \end{cases} \quad (6.46)$$

where $V_\epsilon(t) \equiv \exp_{\bar{y}(t)}^{-1} y^\epsilon(t)$ with $\epsilon > 0$ being small enough. Define

$$\eta(\theta; \epsilon) \equiv \exp_{\bar{y}(t)}(\theta \tilde{V}_\epsilon(t)), \quad \theta \in [0, \rho(\bar{y}(t), y^\epsilon(t))]. \quad (6.47)$$

Then, we have $\eta(0; \epsilon) = \bar{y}(t)$ and $\eta(\rho(\bar{y}(t), y^\epsilon(t)); \epsilon) = y^\epsilon(t)$. For any $h \in C^\infty(M)$, applying Taylor's expansion, Lemma 2.1, (2.17), (5.23) and (5.24), we obtain that

$$\begin{aligned} \frac{\partial}{\partial \epsilon}\bigg|_0 y^\epsilon(t)(h) &= \frac{\partial}{\partial \epsilon}\bigg|_0 h(y^\epsilon(t)) = \lim_{\epsilon \rightarrow 0} \frac{h(y^\epsilon(t)) - h(\bar{y}(t))}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{h(\eta(\rho(\bar{y}(t), y^\epsilon(t)); \epsilon)) - h(\eta(0; \epsilon))}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left(\langle \nabla h(\bar{y}(t)), \frac{\partial}{\partial \theta}\bigg|_0 \eta(0; \epsilon) \rangle \rho(\bar{y}(t), y^\epsilon(t)) + o(\rho(\bar{y}(t), y^\epsilon(t))) \right) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \langle \nabla h(\bar{y}(t)), V_\epsilon(t) \rangle = V(t)(h), \end{aligned}$$

which implies the first equality of (6.45).

To prove the second equality of (6.45), we employ two ways to compute $\frac{\partial^2}{\partial \epsilon^2} \Big|_0 h(y^\epsilon(t))$. On one hand,

$$\begin{aligned} \frac{\partial^2}{\partial \epsilon^2} \Big|_0 h(y^\epsilon(t)) &= \lim_{\epsilon \rightarrow 0} \frac{\partial}{\partial \epsilon} \Big|_0 \langle \nabla h(y^\epsilon(t)), \frac{\partial}{\partial \epsilon} y^\epsilon(t) \rangle \\ &= \langle \nabla_{V(t)} \nabla h, V(t) \rangle + \langle \nabla h(\bar{y}(t)), \frac{\partial^2}{\partial \epsilon^2} \Big|_0 y^\epsilon(t) \rangle, \end{aligned} \quad (6.48)$$

where the first equality of (6.45) is used.

On the other hand, recalling (6.46) and (6.47), we can get, via Taylor's expansion, (2.17) and (5.23),

$$\begin{aligned} &\frac{\partial^2}{\partial \epsilon^2} \Big|_0 h(y^\epsilon(t)) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left(h(y^{2\epsilon}(t)) - 2h(y^\epsilon(t)) + h(\bar{y}(t)) \right) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left\{ h(\eta(\rho(\bar{y}(t), y^{2\epsilon}(t)); 2\epsilon)) - h(\eta(0; 2\epsilon)) - 2 \left[h(\eta(\rho(\bar{y}(t), y^\epsilon(t)); \epsilon)) \right. \right. \\ &\quad \left. \left. - h(\eta(0; \epsilon)) \right] \right\} \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left\{ \langle \nabla h(\bar{y}(t)), \frac{\partial}{\partial \theta} \Big|_0 \eta(\theta; 2\epsilon) \rangle \rho(\bar{y}(t), y^{2\epsilon}(t)) + \frac{1}{2} \frac{\partial^2}{\partial \theta^2} \Big|_0 h(\eta(\theta; 2\epsilon)) \rho^2(\bar{y}(t), y^{2\epsilon}(t)) \right. \\ &\quad \left. - 2 \langle \nabla h(\bar{y}(t)), \frac{\partial}{\partial \theta} \Big|_0 \eta(\theta; \epsilon) \rangle \rho(\bar{y}(t), y^\epsilon(t)) - \frac{\partial^2}{\partial \theta^2} \Big|_0 h(\eta(\theta; \epsilon)) \rho^2(\bar{y}(t), y^\epsilon(t)) + o(\epsilon^2) \right\}. \end{aligned}$$

Since

$$\frac{\partial^2}{\partial \theta^2} \Big|_0 h(\eta(\theta; \epsilon)) = \frac{\partial}{\partial \theta} \Big|_0 \langle \nabla h(\eta(\theta; \epsilon)), \frac{\partial}{\partial \theta} \eta(\theta; \epsilon) \rangle = \frac{1}{\rho^2(\bar{y}(t), y^\epsilon(t))} \langle \nabla_{V_\epsilon(t)} \nabla h, V_\epsilon(t) \rangle,$$

where we have used Lemma 2.1 and the fact that $\beta(\cdot; \epsilon)$ is a geodesic. Thus, by using the above equality and (5.23), we can get

$$\begin{aligned} &\frac{\partial^2}{\partial \epsilon^2} \Big|_0 h(y^\epsilon(t)) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left(\langle \nabla h(\bar{y}(t)), V_{2\epsilon}(t) \rangle + \frac{1}{2} \langle \nabla_{V_{2\epsilon}(t)} \nabla h, V_{2\epsilon}(t) \rangle - 2 \langle \nabla h(\bar{y}(t)), V_\epsilon(t) \rangle \right. \\ &\quad \left. - \langle \nabla_{V_\epsilon(t)} \nabla h, V_\epsilon(t) \rangle + o(\epsilon^2) \right) \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon^2} \left(\langle \nabla h(\bar{y}(t)), 2\epsilon V(t) + 4\epsilon^2 Y(t) \rangle + \frac{1}{2} \langle \nabla_{2\epsilon V(t) + 4\epsilon^2 Y(t)} \nabla h, 2\epsilon V(t) + 4\epsilon^2 Y(t) \rangle \right. \\ &\quad \left. - 2 \langle \nabla h(\bar{y}(t)), \epsilon V(t) + \epsilon^2 Y(t) \rangle - \langle \nabla_{\epsilon V(t) + \epsilon^2 Y(t)} \nabla h, \epsilon V(t) + \epsilon^2 Y(t) \rangle + o(\epsilon^2) \right) \\ &= 2 \langle \nabla h(\bar{y}(t)), Y(t) \rangle + \langle \nabla_{V(t)} \nabla h, V(t) \rangle. \end{aligned}$$

Combining this equality with (6.48), one can get the second equality of (6.45).

Secondly, we claim that $(\nu, \psi_1)^\top \in (-\infty, 0] \times T_{y_1}^* M$ belongs to $(\text{Im} D_{\bar{u}(\cdot)} E)^\perp$ if and only if it satisfies (3.19), where $\bar{\psi}(\cdot)$ is the solution to (3.7) corresponding to the pair (ν, ψ_1) . Thus, by the conditions of Theorem 3.3, the codimension of $\text{Im} D_{\bar{u}(\cdot)} E$ is 1.

In fact, assume that $(\nu, \psi_1) \in (-\infty, 0] \times T_{y_1}^* M$ satisfies (3.19). Recalling (3.4), one can further get

$$\nu \nabla_u f^0[t](v(t)) + \nabla_u f[t](\bar{\psi}(t), v(t)) = 0, \quad a.e. t \in [0, T], \quad \forall v(\cdot) \in L^2(0, T; \mathbb{R}^m). \quad (6.49)$$

Integrating (6.49) over $[0, T]$, we can obtain, via (5.21), (3.7), (6.45) and integration by parts,

$$\begin{aligned}
0 &= \int_0^T \left\{ \nu \nabla_u f^0[t](v(t)) + \nabla_{\dot{y}(t)} V(\bar{\psi}(t)) - \nabla_x f[t](\bar{\psi}(t), V(t)) \right\} dt \\
&= \psi_1(V(T)) + \int_0^T \left\{ \nu \nabla_u f^0[t](v(t)) - \nabla_{\dot{y}(t)} \bar{\psi}(V) - \nabla_x f[t](\bar{\psi}(t), V(t)) \right\} dt \\
&= \nu \int_0^T \left(\nabla_u f^0[t](v(t)) + d_x f^0[t](V(t)) \right) dt + \psi_1(V(T)) \\
&= \nu \frac{d}{d\epsilon} \Big|_0 J(\bar{u}(\cdot) + \epsilon v(\cdot)) + \psi_1 \left(\frac{\partial}{\partial \epsilon} \Big|_0 y^\epsilon(T) \right) \\
&= (\nu, \psi_1)^\top \left(D_{\bar{u}(\cdot)} E(v(\cdot)) \right).
\end{aligned} \tag{6.50}$$

Since $v(\cdot) \in L^2(0, T; \mathbb{R}^m)$ is arbitrarily chosen, we have

$$(\nu, \psi_1)^\top \in (Im D_{\bar{u}(\cdot)} E)^\perp. \tag{6.51}$$

Conversely, if $(\tilde{\nu}, \tilde{\psi}_1)^\top \in (Im D_{\bar{u}(\cdot)} E)^\perp$ with $\tilde{\nu} \leq 0$, one can follow the above argument in an inverse way, and get (3.19) with $\nu = \tilde{\nu}$ and $\bar{\psi}(\cdot) = \tilde{\psi}(\cdot)$, where $\tilde{\psi}(\cdot)$ is the solution to (3.7) with $\nu = \tilde{\nu}$ and $\psi_1 = \tilde{\psi}_1$.

Thirdly, according to (6.44), the above argument and Lemma 6.1, $(\nu, \psi_1)^\top Hess_{\bar{u}(\cdot)} E$ is sign-definite on $Ker D_{\bar{u}(\cdot)} E \times Ker D_{\bar{u}(\cdot)} E$. Now, we need to compute explicitly both $(\nu, \psi_1)^\top Hess_{\bar{u}(\cdot)} E$ and $Ker D_{\bar{u}(\cdot)} E$.

Recalling (6.41), (6.43) and (6.45), we have

$$\begin{aligned}
(\nu, \psi_1)^\top Hess_{\bar{u}(\cdot)} E(v(\cdot), v(\cdot)) &= (\nu, \psi_1)^\top \left(\frac{\partial^2}{\partial \epsilon^2} \Big|_0 E(\bar{u}(\cdot) + \epsilon v(\cdot)) \right) \\
&= \nu \frac{\partial^2}{\partial \epsilon^2} \Big|_0 J(\bar{u}(\cdot) + \epsilon v(\cdot)) + 2\psi_1(Y(T)).
\end{aligned} \tag{6.52}$$

Applying (6.45), we can get

$$\begin{aligned}
&\nu \frac{\partial^2}{\partial \epsilon^2} \Big|_0 J(\bar{u}(\cdot)) \\
&= \nu \frac{\partial}{\partial \epsilon} \Big|_0 \int_0^T \left\{ \langle \nabla_x f^0(t, y^\epsilon(t), \bar{u}(t) + \epsilon v(t)), \frac{\partial}{\partial \epsilon} y^\epsilon(t) \rangle + \nabla_u f^0(t, y^\epsilon(t), \bar{u}(t) + \epsilon v(t))(v(t)) \right\} dt \\
&= \nu \int_0^T \left\{ \nabla_x^2 f^0[t](V(t), V(t)) + 2\nabla_x \nabla_u f^0[t](v(t), V(t)) + \nabla_u^2 f^0[t](v(t), v(t)) \right. \\
&\quad \left. + \langle \nabla_x f^0[t], 2Y(t) \rangle \right\} dt.
\end{aligned} \tag{6.53}$$

Using integration by parts and employing the second order variational equation (5.22), we obtain that

$$\begin{aligned}
&2\psi_1(Y(T)) \\
&= 2 \int_0^T \left[\nabla_{\dot{y}(t)} \bar{\psi}(Y(t)) + \bar{\psi}(t)(\nabla_{\dot{y}(t)} Y) \right] dt \\
&= 2 \int_0^T \left\{ \nabla_{\dot{y}(t)} \bar{\psi}(Y(t)) + \nabla_x f[t](\bar{\psi}(t), Y(t)) + \nabla_u \nabla_x f[t](\bar{\psi}(t), V(t), v(t)) \right. \\
&\quad \left. - \frac{1}{2} R(\bar{\psi}(t), V(t), f[t], V(t)) + \frac{1}{2} \nabla_x^2 f[t](\bar{\psi}(t), V(t), V(t)) + \frac{1}{2} \nabla_u^2 f[t](\bar{\psi}(t), v(t), v(t)) \right\} dt.
\end{aligned}$$

Inserting the above identity and (6.53) into (6.52), one can get, via the first order dual

equation (3.7),

$$\begin{aligned} & (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v(\cdot), v(\cdot)) \\ &= \int_0^T \left\{ \nabla_x^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), V(t)) + \nabla_u^2 H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(v(t), v(t)) \right. \\ & \quad \left. + 2\nabla_u \nabla_x H^\nu(t, \bar{y}(t), \bar{\psi}(t), \bar{u}(t))(V(t), v(t)) - R(\bar{\psi}(t), V(t), f[t], V(t)) \right\} dt. \end{aligned}$$

By the definition of the kernel of $D_{\bar{u}(\cdot)}E$, for any $v(\cdot) \in \text{Ker} D_{\bar{u}(\cdot)}E$, one has

$$0 = \frac{d}{d\epsilon} \Big|_0 \left(\int_0^T f^0(t, y^\epsilon(t), \bar{u}(t) + \epsilon v(t)) dt \right)_{y^\epsilon(T)} = \left(\int_0^T \left\{ \langle \nabla_x f^0[t], V(t) \rangle + \nabla_u f^0[t](v(t)) \right\} dt \right)_{V(T)}. \quad (6.54)$$

In the normal case, i.e. $\nu < 0$, (6.54) implies, together with the first order dual equation (3.7) and (5.21), $V(T) = 0$ and $\int_0^T \nabla_u H^\nu[t](v(t)) dt = 0$. By (3.19), it follows that

$$\text{Ker} D_{\bar{u}(\cdot)}E = \{v(\cdot) \in L^2(0, T; \mathbb{R}^m); \text{ (3.20) with } \xi(\cdot) = v(\cdot) \text{ admits a solution}\}. \quad (6.55)$$

In order to determine the sign of the quadratic $(\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)}E$, from the assumptions, one can find a $u_\eta(\cdot) \in \mathcal{V}_{ad} \setminus \{\bar{u}(\cdot)\}$ such that $\|u_\eta(\cdot) - \bar{u}(\cdot)\|_{L^2(0, T; \mathbb{R}^m)}$ tends to zero as $\eta \rightarrow +\infty$. Set

$$v_\eta(\cdot) = \frac{u_\eta(\cdot) - \bar{u}(\cdot)}{\|u_\eta(\cdot) - \bar{u}(\cdot)\|_{L^2(0, T; \mathbb{R}^m)}}, \quad \epsilon_\eta = \|u_\eta(\cdot) - \bar{u}(\cdot)\|_{L^2(0, T; \mathbb{R}^m)}.$$

Since $D_{\bar{u}(\cdot)}E : L^2(0, T; \mathbb{R}^m) \rightarrow \mathbb{R} \times T_{y_1}M$ is linear, one can find a subspace $N \subset L^2(0, T; \mathbb{R}^m)$ such that $L^2(0, T; \mathbb{R}^m) = \text{Ker} D_{\bar{u}(\cdot)}E \oplus N$, $\text{Im} D_{\bar{u}(\cdot)}E = (D_{\bar{u}(\cdot)}E)N$, and $D_{\bar{u}(\cdot)} : N \rightarrow \mathbb{R} \times T_{y_1}M$ is surjective. Thus, we have $\dim N < \infty$. For each $v_\eta(\cdot)$, there exist a $v_\eta^1(\cdot) \in \text{Ker} D_{\bar{u}(\cdot)}E$ and a $v_\eta^2(\cdot) \in N$ such that $v_\eta(\cdot) = v_\eta^1(\cdot) + v_\eta^2(\cdot)$. In what follows, we will consider two cases.

Case I. If there exists an $n_1 > 0$ such that $v_\eta^2(\cdot) \in N \setminus \{0\}$ for all $n > n_1$, then one can find a subsequence of $\{v_\eta^2(\cdot)\}$ (still denoted by $\{v_\eta^2(\cdot)\}$) and a $v^2(\cdot) \in N$ such that

$$\lim_{\eta \rightarrow +\infty} \|v_\eta^2(\cdot) - v^2(\cdot)\|_{L^2(0, T; \mathbb{R}^m)} = 0. \quad (6.56)$$

We claim that $v^2(\cdot) = 0$.

Indeed, denote by $V_\eta^i(\cdot)$ and $V^2(\cdot)$ the solutions to the first order variational equation (5.21) with $v(\cdot) = v_\eta^i(\cdot)$ for $i = 1, 2$ and $v(\cdot) = v^2(\cdot)$, respectively. Especially, applying (6.56), we have

$$\lim_{\eta \rightarrow +\infty} \max_{t \in [0, T]} |V_\eta^2(t) - V^2(t)| = 0. \quad (6.57)$$

Denote by $V_\eta(\cdot) = V_\eta^1(\cdot) + V_\eta^2(\cdot)$. One can check that $V_\eta(\cdot)$ is the solution to (5.21) with $v(\cdot) = v_\eta(\cdot) = v_\eta^1(\cdot) + v_\eta^2(\cdot)$. Denote by $y_\eta^\epsilon(\cdot)$ the solution to (1.1) with control $u(\cdot) = \bar{u}(\cdot) + \epsilon v_\eta(\cdot)$ for $\epsilon \geq 0$. Set $V_{\eta, \epsilon}(t) = \exp_{\bar{y}(t)}^{-1}(y_\eta^\epsilon(t))$ for $t \in [0, T]$. Applying (5.23), one can obtain

$$V_{\eta, \epsilon}(t) = \epsilon V_\eta(t) + \epsilon^2 Y_\eta(t) + o(\epsilon^2), \quad \forall t \in [0, T],$$

where $Y_\eta(\cdot)$ is the solution to (5.22) with $V(\cdot) = V_\eta(\cdot)$ and $v(\cdot) = v_\eta(\cdot)$. Recalling the proof of Proposition 5.2, one can check that $\epsilon^2 Y_\eta(t) + o(\epsilon^2)$ in the above formula is an infinitesimal $o(\epsilon)$, which is uniform with respect to $v_\eta(\cdot)$, due to $\|v_\eta(\cdot)\|_{L^2(0,T;\mathbb{R}^m)} = 1$. Note that $u_\eta(\cdot) = \bar{u}(\cdot) + \epsilon_\eta v_\eta(\cdot) \in \mathcal{V}_{ad}$, $v_\eta^1(\cdot) \in \text{Ker} D_{\bar{u}(\cdot)} E$ and (6.55). The above formula holds for $\epsilon = \epsilon_\eta$ and

$$0 = V_{\eta,\epsilon_\eta}(T) = \epsilon_\eta V_\eta^2(T) + o(\epsilon_\eta), \quad \eta = n_1 + 1, n_1 + 2, \dots$$

From the above formula and (6.57), one can get $V^2(T) = 0$, and hence $v^2(\cdot) \in \text{Ker} D_{\bar{u}(\cdot)} E \cap N$, which implies the claim.

Applying Taylor's expansion, and noting (6.45) and (6.51) and the boundness of $\{v_\eta(\cdot)\}$ in $L^2(0,T;\mathbb{R}^m)$, one can get

$$\begin{aligned} & \nu \left(J(\bar{u}(\cdot) + \epsilon v_\eta(\cdot)) - J(\bar{u}(\cdot)) \right) + \psi_1(V_{\eta,\epsilon}(T)) \\ &= \epsilon \left(\nu \left(D_{\bar{u}(\cdot)} J \right) v_\eta(\cdot) + \psi_1(V_\eta(T)) \right) + \frac{1}{2} \epsilon^2 \left(\nu \text{Hess}_{\bar{u}(\cdot)} J(v_\eta(\cdot), v_\eta(\cdot)) + \psi_1 \left(\frac{\partial^2}{\partial \epsilon^2} \Big|_0 y_\eta^\epsilon(T) \right) \right) \\ & \quad + o(\epsilon^2) \\ &= \epsilon (\nu, \psi_1)^\top \left(D_{\bar{u}(\cdot)} E v_\eta(\cdot) \right) + \frac{1}{2} \epsilon^2 (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta(\cdot), v_\eta(\cdot)) + o(\epsilon^2) \\ &= \frac{1}{2} \epsilon^2 (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^1(\cdot)) + \epsilon^2 (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^2(\cdot)) \\ & \quad + \frac{1}{2} \epsilon^2 (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^2(\cdot), v_\eta^2(\cdot)) + o(\epsilon^2). \end{aligned} \tag{6.58}$$

Taking $\epsilon = \epsilon_\eta$ in the above formula, we have, via $\nu < 0$, $y_\eta^{\epsilon_\eta}(T) = y_1$, (6.51) and the optimality of $\bar{u}(\cdot)$,

$$\begin{aligned} 0 &\geq \frac{1}{2} (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^1(\cdot)) + (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^2(\cdot)) \\ & \quad + \frac{1}{2} (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^2(\cdot), v_\eta^2(\cdot)) + o(1). \end{aligned} \tag{6.59}$$

By means of the boundedness of $\{v_\eta^2(\cdot)\}$ in $L^2(0,T;\mathbb{R}^m)$, there exists a subsequence of $\{(\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^1(\cdot))\}$ (still denoted by $\{(\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^1(\cdot))\}$) such that $\lim_{\eta \rightarrow +\infty} (\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E(v_\eta^1(\cdot), v_\eta^1(\cdot)) = \xi$ exists. Taking limit in (6.59) as $\eta \rightarrow +\infty$, one can get, via (6.56), $0 \geq \frac{1}{2} \xi$. Since $(\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E$ is sign definite on $\text{Ker} D_{\bar{u}(\cdot)} E$, we get $(\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E$ is negative definite on $\text{Ker} D_{\bar{u}(\cdot)} E$.

Case II. There exists a subsequence of $\{v_\eta(\cdot)\}$ (still denoted by $\{v_\eta(\cdot)\}$) such that $v_\eta(\cdot) \in \text{Ker} D_{\bar{u}(\cdot)} E$. Applying the same argument as that in (6.58) and (6.59), one can get $(\nu, \psi_1)^\top \text{Hess}_{\bar{u}(\cdot)} E$ is negative definite on $\text{Ker} D_{\bar{u}(\cdot)} E$.

In the abnormal case, i.e. $\nu = 0$, $\text{Ker} D_{\bar{u}(\cdot)} E$ is equal to (3.22). Lemma 6.1 concludes that the left hand of (3.21) is sign-definite on (3.22). \square

Finally, let us prove Theorem 3.4.

Proof of Theorem 3.4 Theorem 3.4 is a direct consequence of [1, Theorem 21.8, p. 347]. \square

References

- [1] A. A. Agrachev and Y. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Encyclopaedia of Mathematical Sciences, vol. 87, Springer-Verlag, Berlin, 2004.
- [2] D. J. Bell and D. H. Jacobson, *Singular Optimal Control Problems*, Academic Press, London-New York, 1975.
- [3] J. F. Bonnans and A. Hermant, Second-order Analysis for Optimal Control Problems with Pure State Constraints and Mixed Control-State Constraints, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26 (2009), 561–598.
- [4] J. F. Bonnans and A. Hermant, No-gap Second-order Optimality Conditions for Optimal Control Problems with a Single State Constraint and Control, *Math. Program.*, Ser. B, 117 (2009) 21–50.
- [5] B. Bonnard, J. B. Caillau and E. Trélat, Second Order Optimality Conditions in the Smooth Case and Applications in Optimal Control, *ESAIM Control Optim. Calc. Var.*, 13 (2007), 207–236.
- [6] B. Bonnard and M. Chyba, *Singular Trajectories and Their Role in Control Theory*, Mathématiques & Applications, vol. 40, Springer-Verlag, Berlin, Heidelberg, New York, 2003.
- [7] U. Boschain and B. Piccoli, *Optimal Syntheses for Control Systems on 2-D Manifolds*, Mathématiques & Applications, vol. 43, Springer-Verlag, Berlin, Heidelberg, New York, 2004.
- [8] D. J. Clements and B. D. O. Anderson, *Singular Optimal Control: the Linear-Quadratic Problem*, Springer-Verlag, Berlin, New York, 1978.
- [9] Q. Cui, L. Deng and X. Zhang, Pointwise Second Order Necessary Conditions for Optimal Control Problems Evolved on Riemannian Manifolds, *C. R. Math. Acad. Sci. Paris, Ser. I*, 354 (2016), 191–194.
- [10] L. Deng, Dynamic Programming Method for Control Systems on Manifolds and its Relations to Maximum Principle, *J. Math. Anal. Appl.*, 434 (2016), 915–938.
- [11] M. P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1976.
- [12] M. P. do Carmo, *Riemannian Geometry*, Translated from the Second Portuguese Edition by F. Flaherty, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1992.
- [13] H. Frankowska and D. Tonon, Pointwise Second-order Necessary Optimality Conditions for the Mayer Problem with Control Constraints, *SIAM J. Control Optim.*, 51 (2013), 3814–3843.
- [14] R. F. Gabasov and F. M. Kirillova, High Order Necessary Conditions for Optimality, *SIAM J. Control*, 10 (1972), 127–168.

- [15] R. F. Gabasov and F. M. Kirillova, *Singular Optimal Controls*, Izdat. "Nauka", Moscow, 1973.
- [16] B. S. Goh, The Second Variation for the Singular Bolza Problem, *SIAM J. Control*, 4 (1966), 309–325.
- [17] B. S. Goh, Necessary Conditions for Singular Extremals Involving Multiple Control Variables, *SIAM J. Control*, 4 (1966), 716–731.
- [18] S. Helgason, *Differential Geometry and Symmetric Spaces*, Academic press, New York and London, 1962.
- [19] D. Hoehener, Variational Approach to Second-order Optimality Conditions for Control Problems with Pure State Constraints, *SIAM. J. Control Optim.*, 50 (2012), 1139–1173.
- [20] D. Hoehener, Variational Approach to Second-order Sufficient Optimality Conditions in Optimal Control, *SIAM. J. Control Optim.*, 52 (2014), 861–892.
- [21] H.-W. Knobloch, *Higher Order Necessary Conditions in Optimal Control Theory*, Springer-Verlag, Berlin-New York, 1981.
- [22] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, vol.1, Interscience, New York London, 1963.
- [23] A. J. Krener, The High Order Maximal Principle and its Application to Singular Extremals, *SIAM J. Control Optim.*, 15 (1977), 256–293.
- [24] J. M. Lee, *Riemannian Manifolds: An Introduction to Curvature*, Graduate Texts in Mathematics, vol. 176, Springer-Verlag, New York, 1997.
- [25] X. Li and J. Yong, *Optimal Control Theory for Infinite-Dimensional Systems*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 1995.
- [26] H. Lou, Second-order Necessary/Sufficient Conditions for Optimal Control Problems in the Absence of Linear Structure, *Discrete Contin. Dyn. Syst. Ser. B*, 14 (2010), 1445–1464.
- [27] K. Malanowski, Sufficient Optimality Conditions for Optimal Control Subject to State Constraints, *SIAM J. Control Optim.*, 35 (1997), 205–227.
- [28] K. Malanowski, H. Maurer and S. Pickenhain, Second-order Sufficient Conditions for State-constrained Optimal Control Problems, *J. Optim. Theory Appl.*, 123 (2004), 595–617.
- [29] N. P. Osmolovskii and H. Maurer, *Applications to Regular and Bang-Bang Control. Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control*, SIAM, Philadelphia, PA, 2012.
- [30] Z. Páles and V. Zeidan, Optimal Control Problems with Set-valued Control and State Constraints, *SIAM J. Optim.*, 14 (2003), 334–358.

- [31] P. Petersen, *Riemannian Geometry*, Second Edition, Graduate Texts in Mathematics, vol. 171, Springer-Verlag, New York, 2006.
- [32] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mischenko, *Mathematical Theory of Optimal Processes*, Wiley, New York, 1962.
- [33] H. Schättler and U. Ledzewicz, *Geometric Optimal Control, Theory, Methods and Examples*, Interdisciplinary Applied Mathematics, vol. 38, Springer, New York, 2012.
- [34] H. J. Sussmann, Geometry and Optimal Control, in: *Mathematical Control Theory, Calculus of Variations and Optimal Control, Optimization*, Springer, New York, 1999, 140–198.
- [35] J. Warga, A Second-order Condition that Strengthens Pontryagin’s Maximum Principle, *J. Differential Equations*, 28 (1978), 284–307.
- [36] H. Wu, C. L. Shen and Y. L. Yu, *An Introduction to Riemannian Geometry* (In Chinese), Press of Peking University, Beijing, 1989.